(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 January 2003 (16.01.2003)

PCT

(10) International Publication Number
WO 03/004629 A2

(51) International Patent Classification[7]: C12N 9/00

(21) International Application Number: PCT/EP02/07229

(22) International Filing Date: 1 July 2002 (01.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

| | | |
|---|---|---|
| 60/301,853 | 2 July 2001 (02.07.2001) | US |
| 60/337,130 | 10 December 2001 (10.12.2001) | US |
| 60/375,015 | 25 April 2002 (25.04.2002) | US |

(71) Applicant *(for all designated States except US)*: **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **ZHU, Zhimin** [CN/US]; 45 Hinckley Road, Waban, MA 02468 (US).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declarations under Rule 4.17:**

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

**Published:**

— *without international search report and to be republished upon receipt of that report*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 03/004629 A2

(54) Title: REGULATION OF HUMAN CITRON RHO/RAC-INTERACTING KINASE-SHORT KINASE

(57) Abstract: Reagents that regulate human CRIK-sk and reagents which bind to human CRIK-sk gene products can play a role in preventing, ameliorating, or correcting dysfunctions or diseases including, but not limited to, obesity, diabetes, cancer or COPD.

# REGULATION OF HUMAN CITRON RHO/RAC-INTERACTING KINASE-SHORT KINASE

This application incorporates by reference and claims the benefit of co-pending provisional applications Serial No. 60/301,853 filed July 2, 2001, Serial No. 60/337,130 filed December 10, 2001, and Serial Number 60/375,015 filed April 25, 2002.

## TECHNICAL FIELD OF THE INVENTION

The invention relates to the regulation of human citron rho/rac-interacting kinase-short kinase (CRIK-sk).

## BACKGROUND OF THE INVENTION

Kinases are involved in a variety of disease processes. There is a need in the art to identify related enzymes, which can be regulated for therapeutic effects.

## SUMMARY OF THE INVENTION

It is an object of the invention to provide reagents and methods of regulating a human CRIK-sk. This and other objects of the invention are provided by one or more of the embodiments described below.

One embodiment of the invention is a human citron rho/rac-interacting kinase-short kinase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 88% identical to
the amino acid sequence shown in SEQ ID NO: 2;
the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 88% identical to the amino acid sequence shown in SEQ ID NO: 9; and
the amino acid sequence shown in SEQ ID NO: 9.

Yet another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a human citron rho/rac-interacting kinase-short kinase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 88% identical to the amino acid sequence shown in SEQ ID NO: 2;
the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 88% identical to the amino acid sequence shown in SEQ ID NO: 9; and
the amino acid sequence shown in SEQ ID NO: 9.

Binding between the test compound and the human citron rho/rac-interacting kinase-short kinase polypeptide is detected. A test compound which binds to the human citron rho/rac-interacting kinase-short kinase polypeptide is thereby identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the activity of the human citron rho/rac-interacting kinase-short kinase.

Another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a polynucleotide encoding a human citron rho/rac-interacting kinase-short kinase polypeptide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;
the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 8; and
the nucleotide sequence shown in SEQ ID NO: 8.

Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the amount of the human citron rho/rac-interacting kinase-short kinase through interacting with the human citron rho/rac-interacting kinase-short kinase mRNA.

Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a human citron rho/rac-interacting kinase-short kinase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 88% identical to the amino acid sequence shown in SEQ ID NO: 2;
the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 88% identical to the amino acid sequence shown in SEQ ID NO: 9; and
the amino acid sequence shown in SEQ ID NO: 9.

A human citron rho/rac-interacting kinase-short kinase activity of the polypeptide is detected. A test compound which increases human citron rho/rac-interacting kinase-short kinase activity of the polypeptide relative to human citron rho/rac-interacting kinase-short kinase activity in the absence of the test compound is thereby identified

as a potential agent for increasing extracellular matrix degradation. A test compound which decreases human citron rho/rac-interacting kinase-short kinase activity of the polypeptide relative to human citron rho/rac-interacting kinase-short kinase activity in the absence of the test compound is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Even another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a human citron rho/rac-interacting kinase-short kinase product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;
the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 8; and
the nucleotide sequence shown in SEQ ID NO: 8.

Binding of the test compound to the human citron rho/rac-interacting kinase-short kinase product is detected. A test compound which binds to the human citron rho/rac-interacting kinase-short kinase product is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a human citron rho/rac-interacting kinase-short kinase polypeptide or the product encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to
the nucleotide sequence shown in SEQ ID NO: 1;
the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to
the nucleotide sequence shown in SEQ ID NO: 8; and
the nucleotide sequence shown in SEQ ID NO: 8.

Human citron rho/rac-interacting kinase-short kinase activity in the cell is thereby decreased.

The invention thus provides a human CRIK-sk that can be used to identify test compounds that may act, for example, as activators or inhibitors at the enzyme's active site. Human CRIK-sk and fragments thereof also are useful in raising specific antibodies that can block the enzyme and effectively reduce its activity.

## BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the DNA-sequence encoding a human citron rho/rac-interacting kinase-short kinase Polypeptide (SEQ ID NO: 1).

Fig. 2 shows the amino acid sequence deduced from the DNA-sequence of Fig.1 (SEQ ID NO: 2).

Fig. 3 shows the amino acid sequence of the protein identified by trembl|AF086823|AF086823_1 (SEQ ID NO: 3).

Fig. 4 shows the amino acid sequence of the protein identified by trembl|AF086824|AF086824_1 (SEQ ID NO: 4).

Fig. 5 shows the amino acid sequence of the protein identified by trembl|AF128625|AF128625_1 (SEQ ID NO: 5).

Fig. 6 shows the DNA-sequence encoding a human citron rho/rac-interacting kinase-short kinase Polypeptide (SEQ ID NO: 6).

Fig. 7 shows the amino acid sequence of the protein identified by swissnew|P54265|DMK_MOUSE (SEQ ID NO: 7)

Fig. 8 shows the BLASTP - alignment of CRIK-sk (SEQ ID NO: 2) against trembl|AF086823|AF086823_1 (SEQ ID NO: 3).

Fig. 9 shows the BLASTP - alignment of CRIK-sk (SEQ ID NO: 2) against trembl|AF086824|AF086824_1 (SEQ ID NO: 4).

Fig. 10 shows the BLASTP - alignment of CRIK-sk (SEQ ID NO: 2) against trembl|AF128625|AF128625_1 (SEQ ID NO: 5).

Fig. 11 shows the BLASTP - alignment of CRIK-sk (SEQ ID NO: 2) against swissnew|P54265|DMK_MOUSE.

Fig. 12 shows the BLASTP - alignment of CRIK-sk (SEQ ID NO: 2) against pdb|1CDK|1CDK-A.

Fig. 13 shows the HMMPFAM - alignment of CRIK-sk (SEQ ID NO: 2) against pfam|hmm|pkinase

Fig. 14 shows the HMMPFAM - alignment of CRIK-sk (SEQ ID NO: 2) against pfam|hmm|pkinase_C

Fig. 15 shows the Prosite search results.

Fig. 16 shows the Genewise output.

Fig. 17 shows the Relative expression of human citron rho/rac-interacting kinase-short kinase.

Fig. 18 shows the shows the DNA-sequence encoding a human citron rho/rac-interacting kinase-short kinase Polypeptide (SEQ ID NO: 8)

Fig. 19 shows the amino acid sequence deduced from the DNA-sequence of Fig. 18 (SEQ ID NO: 9)

Fig. 20 shows the TBLASTN – alignment of 544_Protein against emnew|AX166510|AX166510 Sequence 1 from Patent WO0138503.//:gbnew|AX166510|AX66510 Sequence Patent WO0138503

## DETAILED DESCRIPTION OF THE INVENTION

The invention relates to an isolated polynucleotide from the group consisting of:

a) a polynucleotide encoding a human citron rho/rac-interacting kinase-short kinase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 88% identical to the amino acid sequence shown in SEQ ID NO: 2;
the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 88% identical to the amino acid sequence shown in SEQ ID NO: 9; and
the amino acid sequence shown in SEQ ID NO: 9.

b) a polynucleotide comprising the sequence of SEQ ID NOS: 1 or 8;

c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b) and encodes a human citron rho/rac-interacting kinase-short kinase polypeptide;

d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code and encodes a human citron rho/rac-interacting kinase-short kinase polypeptide; and

e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d) and encodes a human citron rho/rac-interacting kinase-short kinase polypeptide.

Furthermore, it has been discovered by the present applicant that a novel CRIK-sk, particularly a human CRIK-sk which is differentially expressed in the hypothalamus, can be used in therapeutic methods to treat obesity, diabetes, cancer or COPD. Human CRIK-sk comprises the amino acid sequence shown in SEQ ID NO: 2. A coding sequence for human CRIK-sk is shown in SEQ ID NO: 1. This sequence is located on chromosome 12q24.2. Related ESTs are expressed in uterus_tumor, glioblastoma with EGFR amplification, colon, and normal nervous tissue.

Human CRIK-sk is 87% identical over 495 amino acids to trembl|AF086823|AF086823_1 (SEQ ID NO: 3) (FIG. 1), 88% identical over 468 amino acids to trembl|AF086824|AF086824_1 (SEQ ID NO: 4) (FIG. 2), 42% identical over 420 amino acids to trembl|AF128625|AF128625_1 (SEQ ID NO: 5) (FIG. 3), 44% identical over 386 amino acids to swissnew|P54265|DMK_MOUSE (SEQ ID NO: 11) (FIG. 4), and 33% identical over 333 amino acids to pdb|1CDK|1CDK-A (FIG. 5).

Human CRIK-sk of the invention is expected to be useful for the same purposes as previously identified CRIK-sk enzymes. Human CRIK-sk is believed to be useful in therapeutic methods to treat disorders such as obesity and COPD. Human CRIK-sk also can be used to screen for human CRIK-sk activators and inhibitors.

*Polypeptides*

Human CRIK-sk polypeptides according to the invention comprise at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, or 495 contiguous amino acids selected from the amino acid sequence shown in SEQ ID NO: 2 or a biologically active variant thereof, as defined below. A CRIK-sk polypeptide of the invention therefore can be a portion of a CRIK-sk protein, a full-length CRIK-sk protein, or a fusion protein comprising all or a portion of a CRIK-sk protein.

*Biologically Active Variants*

Human CRIK-sk polypeptide variants which are biologically active, *e.g.*, retain enzymatic activity, also are human CRIK-sk polypeptides. Preferably, naturally or non-naturally occurring human CRIK-sk polypeptide variants have amino acid sequences which are at least about 88, 90, 96, or 98 or 99% identical to the amino acid sequence shown in SEQ ID NO: 2 or a fragment thereof. Percent identity between a putative human CRIK-sk polypeptide variant and an amino acid sequence of SEQ ID NO: 2 is determined by conventional methods. See, for example, Altschul *et al.*, *Bull. Math. Bio. 48*:603 (1986), and Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA 89*:10915 (1992). Briefly, two amino acid sequences are aligned to optimize the alignment scores using a gap opening penalty of 10, a gap extension penalty of 1, and the "BLOSUM62" scoring matrix of Henikoff & Henikoff, 1992.

Those skilled in the art appreciate that there are many established algorithms available to align two amino acid sequences. The "FASTA" similarity search algorithm of Pearson & Lipman is a suitable protein alignment method for examining the level of identity shared by an amino acid sequence disclosed herein and the amino acid sequence of a putative variant. The FASTA algorithm is described by Pearson & Lipman, *Proc. Nat'l Acad. Sci. USA 85*:2444(1988), and by Pearson, *Meth. Enzymol. 183*:63 (1990). Briefly, FASTA first characterizes sequence similarity by identifying regions shared by the query sequence (*e.g.*, SEQ ID NO: 2) and a test sequence that have either the highest density of identities (if the ktup variable is 1) or pairs of identities (if ktup=2), without considering conservative amino acid substitutions, insertions, or deletions. The ten regions with the highest density of identities are then rescored by comparing the similarity of all paired amino acids using an amino acid substitution matrix, and the ends of the regions are "trimmed" to include only those residues that contribute to the highest score. If there are several regions with scores greater than the "cutoff" value (calculated by a predetermined formula based upon the length of the sequence the ktup value), then the trimmed initial regions are examined to determine whether the regions can be joined to form an approximate alignment with gaps. Finally, the highest scoring regions of the two amino acid sequences are aligned using a modification of the Needleman-Wunsch-Sellers algorithm (Needleman & Wunsch, *J. Mol. Biol.48*:444 (1970); Sellers, *SIAM J. Appl. Math.26*:787 (1974)), which allows for amino acid insertions and deletions. Preferred parameters for FASTA analysis are: ktup=1, gap opening penalty=10, gap extension penalty=1, and substitution matrix=BLOSUM62. These parameters can be introduced into a FASTA program by modifying the scoring matrix file ("SMATRIX"), as explained in Appendix 2 of Pearson, *Meth. Enzymol. 183*:63 (1990).

FASTA can also be used to determine the sequence identity of nucleic acid molecules using a ratio as disclosed above. For nucleotide sequence comparisons, the ktup value can range between one to six, preferably from three to six, most preferably three, with other parameters set as default.

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a human CRIK-sk polypeptide can be found using computer programs well known in the art, such as DNASTAR software.

The invention additionally, encompasses CRIK-sk polypeptides that are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications can be carried out by known techniques including, but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, $NaBH_4$, acetylation, formylation, oxidation, reduction, metabolic synthesis in the presence of tunicamycin, etc.

Additional post-translational modifications encompassed by the invention include, for example, *e.g.*, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The CRIK-sk polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

The invention also provides chemically modified derivatives of CRIK-sk polypeptides that may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent No. 4,179,337). The chemical moieties for derivitization can be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol, and the like. The polypeptides can be modified at random or predetermined positions within the molecule and can include one, two, three, or more attached chemical moieties.

Whether an amino acid change or a polypeptide modification results in a biologically active CRIK-sk polypeptide can readily be determined by assaying for enzymatic activity, as described for example, in Di Cunto F. *et al.*, J Biol Chem. 1998 Nov 6;273(45):29706-11.

*Fusion Proteins*

Fusion proteins are useful for generating antibodies against CRIK-sk polypeptide amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins that interact with portions of a CRIK-sk polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A CRIK-sk polypeptide fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, or 495 contiguous amino acids of SEQ ID NO: 2 or of a biologically active variant, such as those described above. The first polypeptide segment also can comprise full-length CRIK-sk protein.

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the CRIK-sk polypeptide-encoding sequence and the heterologous protein sequence, so that the CRIK-sk polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises coding sequences selected from SEQ ID NO: 1 in proper reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

*Identification of Species Homologs*

Species homologs of human CRIK-sk polypeptide can be obtained using CRIK-sk polypeptide polynucleotides (described below) to make suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologs of CRIK-sk polypeptide, and expressing the cDNAs as is known in the art.

*Polynucleotides*

A CRIK-sk polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a CRIK-sk polypeptide. A coding sequence for human CRIK-sk is shown in SEQ ID NO: 1.

Degenerate nucleotide sequences encoding human CRIK-sk polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, 98, or 99% identical to the nucleotide sequence shown in SEQ ID NO: 1 or its complement also are CRIK-sk polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of CRIK-sk polynucleotides that encode biologically active CRIK-sk polypeptides also are CRIK-sk polynucleotides. Polynucleotide fragments comprising at least 8, 9, 10, 11, 12, 15, 20, or 25 contiguous nucleotides of SEQ ID NO: 1 or its complement also are CRIK-sk polynucleotides. These fragments can be used, for example, as hybridization probes or as antisense oligonucleotides.

*Identification of Polynucleotide Variants and Homologs*

Variants and homologs of the CRIK-sk polynucleotides described above also are CRIK-sk polynucleotides. Typically, homologous CRIK-sk polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known CRIK-sk polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions--2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1% SDS, 50°C once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each--homologous sequences can be identified which contain at most about 25-30% base pair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologs of the CRIK-sk polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of CRIK-sk polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the $T_m$ of a double-stranded DNA decreases by 1-1.5°C with every 1% decrease in homology (Bonner *et al.*, *J. Mol. Biol. 81*, 123 (1973). Variants of human CRIK-sk polynucleotides or CRIK-sk polynucleotides of other species can therefore be identified by hybridizing a putative homologous CRIK-sk polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO: 1 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to CRIK-sk polynucleotides or their complements following stringent hybridization and/or wash conditions also are CRIK-sk

polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20°C below the calculated $T_m$ of the hybrid under study. The $T_m$ of a hybrid between a CRIK-sk polynucleotide having a nucleotide sequence shown in SEQ ID NO: 1 or the complement thereof and a polynucleotide sequence which is at least about 50, preferably about 75, 90, 96, or 98% identical to one of those nucleotide sequences can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A. 48*, 1390 (1962):

$$T_m = 81.5°C - 16.6(\log_{10}[Na^+]) + 0.41(\%G + C) - 0.63(\%\text{formamide}) - 600/l),$$

where $l$ = the length of the hybrid in base pairs.

Stringent wash conditions include, for example, 4X SSC at 65°C, or 50% formamide, 4X SSC at 42°C, or 0.5X SSC, 0.1% SDS at 65°C. Highly stringent wash conditions include, for example, 0.2X SSC at 65°C.

*Preparation of Polynucleotides*

A CRIK-sk polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, or synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or by using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated CRIK-sk polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments, which comprise CRIK-sk

nucleotide sequences. Isolated polynucleotides are in preparations that are free or at least 70, 80, or 90% free of other molecules.

Human CRIK-sk cDNA molecules can be made with standard molecular biology techniques, using CRIK-sk mRNA as a template. Human CRIK-sk cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook *et al.* (1989). An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesize CRIK-sk polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a CRIK-sk polypeptide having, for example, an amino acid sequence shown in SEQ ID NO: 2 or a biologically active variant thereof.

*Extending Polynucleotides*

Various PCR-based methods can be used to extend the nucleic acid sequences disclosed herein to detect upstream sequences such as promoters and regulatory elements. For example, restriction-site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic. 2*, 318-322, 1993). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al.*, *Nucleic Acids Res. 16*, 8186, 1988). Primers

can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72°C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al.*, *PCR Methods Applic. 1*, 111-119, 1991). In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which can be used to retrieve unknown sequences is that of Parker *et al.*, *Nucleic Acids Res. 19*, 3055-3060, 1991). Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA (CLONTECH, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Randomly-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' non-transcribed regulatory regions.

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) that are laser

activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (*e.g.* GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA that might be present in limited amounts in a particular sample.

*Obtaining Polypeptides*

Human CRIK-sk polypeptides can be obtained, for example, by purification from human cells, by expression of CRIK-sk polynucleotides, or by direct chemical synthesis.

*Protein Purification*

Human CRIK-sk polypeptides can be purified from any cell that expresses the polypeptide, including host cells that have been transfected with CRIK-sk expression constructs. A purified CRIK-sk polypeptide is separated from other compounds that normally associate with the CRIK-sk polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified CRIK-sk polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

*Expression of Polynucleotides*

To express a CRIK-sk polynucleotide, the polynucleotide can be inserted into an expression vector that contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods that are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding CRIK-sk polypeptides and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et al.* (1989) and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

A variety of expression vector/host systems can be utilized to contain and express sequences encoding a CRIK-sk polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those non-translated regions of the vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*,

viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a CRIK-sk polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

*Bacterial and Yeast Expression Systems*

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the CRIK-sk polypeptide. For example, when a large quantity of a CRIK-sk polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene). In a BLUESCRIPT vector, a sequence encoding the CRIK-sk polypeptide can be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β-galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem. 264*, 5503-5509, 1989) or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel *et al.* (1989) and Grant *et al.*, *Methods Enzymol. 153*, 516-544, 1987.

*Plant and Insect Expression Systems*

If plant expression vectors are used, the expression of sequences encoding CRIK-sk polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be used alone or in combination with the omega leader sequence from TMV (Takamatsu, *EMBO J. 6*, 307-311, 1987). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (Coruzzi *et al.*, *EMBO J. 3*, 1671-1680, 1984; Broglie *et al.*, *Science 224*, 838-843, 1984; Winter *et al.*, *Results Probl. Cell Differ. 17*, 85-105, 1991). These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (*e.g.*, Hobbs or Murray, in MCGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill, New York, N.Y., pp. 191-196, 1992).

An insect system also can be used to express a CRIK-sk polypeptide. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. Sequences encoding CRIK-sk polypeptides can be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of CRIK-sk polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect *S. frugiperda cells* or *Trichoplusia* larvae in which CRIK-sk polypeptides can be expressed (Engelhard *et al.*, *Proc. Nat. Acad. Sci. 91*, 3224-3227, 1994).

*Mammalian Expression Systems*

A number of viral-based expression systems can be used to express CRIK-sk polypeptides in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding CRIK-sk polypeptides can be ligated into an

adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome can be used to obtain a viable virus that is capable of expressing a CRIK-sk polypeptide in infected host cells (Logan & Shenk, *Proc. Natl. Acad. Sci. 81*, 3655-3659, 1984). If desired, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (*e.g.*, liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding CRIK-sk polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a CRIK-sk polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used (see Scharf *et al.*, *Results Probl. Cell Differ. 20*, 125-162, 1994).

*Host Cells*

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed CRIK-sk polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to,

acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells that have specific cellular machinery and characteristic mechanisms for post-translational activities (*e.g.*, CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express CRIK-sk polypeptides can be transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced CRIK-sk sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. See, for example, ANIMAL CELL CULTURE, R.I. Freshney, ed., 1986.

Any number of selection systems can be used to recover transformed cell lines.

These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell 11*, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy *et al.*, *Cell 22*, 817-23, 1980) genes which can be employed in *tk*$^-$ or *aprt*$^-$ cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci. 77*, 3567-70, 1980), *npt* confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol. 150*, 1-14, 1981), and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin

acetyltransferase, respectively (Murray, 1992, *supra*). Additional selectable genes have been described. For example, *trpB* allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci. 85*, 8047-51, 1988). Visible markers such as anthocyanins, β-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al.*, *Methods Mol. Biol. 55*, 121-131, 1995).

*Detecting Expression*

Although the presence of marker gene expression suggests that the CRIK-sk polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding a CRIK-sk polypeptide is inserted within a marker gene sequence, transformed cells containing sequences that encode a CRIK-sk polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a CRIK-sk polypeptide under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the CRIK-sk polynucleotide.

Alternatively, host cells which contain a CRIK-sk polynucleotide and which express a CRIK-sk polypeptide can be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques that include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein. For example, the presence of a polynucleotide sequence encoding a CRIK-sk polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a CRIK-sk polypeptide. Nucleic acid amplification-

based assays involve the use of oligonucleotides selected from sequences encoding a CRIK-sk polypeptide to detect transformants that contain a CRIK-sk polynucleotide.

A variety of protocols for detecting and measuring the expression of a CRIK-sk polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a CRIK-sk polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med. 158*, 1211-1216, 1983).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding CRIK-sk polypeptides include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a CRIK-sk polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

*Expression and Purification of Polypeptides*

Host cells transformed with nucleotide sequences encoding a CRIK-sk polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode CRIK-sk polypeptides can be designed to contain signal sequences which direct secretion of soluble CRIK-sk polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound CRIK-sk polypeptide.

As discussed above, other constructions can be used to join a sequence encoding a CRIK-sk polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the CRIK-sk polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a CRIK-sk polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography, as described in Porath *et al.*, *Prot. Exp. Purif. 3*, 263-281, 1992), while the enterokinase cleavage site provides a means for purifying the CRIK-sk polypeptide from the fusion protein. Vectors that contain fusion proteins are disclosed in Kroll *et al.*, *DNA Cell Biol. 12*, 441-453, 1993.

*Chemical Synthesis*

Sequences encoding a CRIK-sk polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*, *Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al. Nucl. Acids Res. Symp. Ser.* 225-232, 1980). Alternatively, a CRIK-sk polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc. 85*, 2149-2154, 1963; Roberge *et al.*, *Science 269*, 202-204, 1995). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Optionally, fragments of CRIK-sk polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, WH Freeman and Co., New York, N.Y., 1983). The composition of a synthetic CRIK-sk polypeptide can be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure; *see* Creighton, *supra*). Additionally, any portion of the amino acid sequence of the CRIK-sk polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

*Production of Altered Polypeptides*

As will be understood by those of skill in the art, it may be advantageous to produce CRIK-sk polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to

produce an RNA transcript having desirable properties, such as a half-life that is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter CRIK-sk polypeptide-encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

*Antibodies*

Any type of antibody known in the art can be generated to bind specifically to an epitope of a CRIK-sk polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv, which are capable of binding an epitope of a CRIK-sk polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, *e.g.*, at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of a CRIK-sk polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody that specifically binds to the immunogen.

Typically, an antibody which specifically binds to a CRIK-sk polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to CRIK-sk polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a CRIK-sk polypeptide from solution.

Human CRIK-sk polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a CRIK-sk polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (*e.g.*, aluminum hydroxide), and surface active substances (*e.g.* lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (*bacilli Calmette-Guerin*) and *Corynebacterium parvum* are especially useful.

Monoclonal antibodies that specifically bind to a CRIK-sk polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al.*, *Nature 256*, 495-497, 1985; Kozbor *et al.*, *J. Immunol. Methods 81*, 31-42, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci. 80*, 2026-2030, 1983; Cole *et al.*, *Mol. Cell Biol. 62*, 109-120, 1984).

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*, *Proc. Natl. Acad. Sci. 81*, 6851-6855, 1984; Neuberger *et al.*, *Nature 312*, 604-608,

1984; Takeda *et al.*, *Nature 314*, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, humanized antibodies can be produced using recombinant methods, as described in GB2188638B. Antibodies that specifically bind to a CRIK-sk polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies that specifically bind to CRIK-sk polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobin libraries (Burton, *Proc. Natl. Acad. Sci. 88*, 11120-23, 1991).

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev. 5*, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol. 15*, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem. 269*, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding

sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology (Verhaar *et al.*, 1995, *Int. J. Cancer 61*, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth. 165*, 81-91):

Antibodies which specifically bind to CRIK-sk polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci. 86*, 3833-3837, 1989; Winter *et al.*, *Nature 349*, 293-299, 1991).

Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

Antibodies according to the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to which a CRIK-sk polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

*Antisense Oligonucleotides*

Antisense oligonucleotides are nucleotide sequences that are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of CRIK-sk gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters. *See* Brown, *Meth. Mol. Biol. 20*, 1-8, 1994; Sonveaux, *Meth. Mol. Biol. 26*, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev. 90*, 543-583, 1990.

Modifications of CRIK-sk gene expression can be obtained by designing antisense oligonucleotides that will form duplexes to the control, 5', or regulatory regions of the CRIK-sk gene. Oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (*e.g.*, Gee *et al.*, in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Futura Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful complex formation between an antisense oligonucleotide and the complementary sequence of a CRIK-sk polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a CRIK-sk polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent CRIK-sk nucleotides, can provide sufficient targeting specificity for CRIK-sk mRNA. Preferably, each stretch of complementary

contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular CRIK-sk poly-nucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to a CRIK-sk polynucleotide. These modifications can be internal or at one or both ends of the antisense molecule. For example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5'-substituted oligonucleotide in which the 3' hydroxyl group or the 5' phosphate group are substituted, also can be employed in a modified antisense oligonucleotide. These modified oligonucleotides can be prepared by methods well known in the art. *See, e.g.*, Agrawal *et al.*, *Trends Biotechnol. 10*, 152-158, 1992; Uhlmann *et al.*, *Chem. Rev. 90*, 543-584, 1990; Uhlmann *et al.*, *Tetrahedron. Lett. 215*, 3539-3542, 1987.

*Ribozymes*

Ribozymes are RNA molecules with catalytic activity. *See, e.g.*, Cech, *Science 236*, 1532-1539; 1987; Cech, *Ann. Rev. Biochem. 59*, 543-568; 1990, Cech, *Curr. Opin. Struct. Biol. 2*, 605-609; 1992, Couture & Stinchcomb, *Trends Genet. 12*, 510-515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (*e.g.*, Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of a CRIK-sk polynucleotide can be used to generate ribozymes that will specifically bind to mRNA transcribed from the CRIK-sk polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (*see* Haseloff *et al. Nature 334*, 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).

Specific ribozyme cleavage sites within a CRIK-sk RNA target can be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate CRIK-sk RNA targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease CRIK-sk expression. Alternatively, if it is desired that the cells stably retain the DNA construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A ribozyme-encoding

DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors that induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

*Differentially Expressed Genes*

Described herein are methods for the identification of genes whose products interact with human CRIK-sk. Such genes may represent genes that are differentially expressed in disorders including, but not limited to, obesity and COPD.

Further, such genes may represent genes that are differentially regulated in response to manipulations relevant to the progression or treatment of such diseases. Additionally, such genes may have a temporally modulated expression, increased or decreased at different stages of tissue or organism development. A differentially expressed gene may also have its expression modulated under control versus experimental conditions. In addition, the human CRIK-sk gene or gene product may itself be tested for differential expression.

The degree to which expression differs in a normal versus a diseased state need only be large enough to be visualized via standard characterization techniques such as differential display techniques. Other such standard characterization techniques by which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase), PCR, and Northern analysis.

*Identification of Differentially Expressed Genes*

To identify differentially expressed genes total RNA or, preferably, mRNA is isolated from tissues of interest. For example, RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique that does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Ausubel *et al.*, ed., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

Transcripts within the collected RNA samples that represent RNA produced by differentially expressed genes are identified by methods well known to those of skill in the art. They include, for example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 85*, 208-12, 1988), subtractive hybridization (Hedrick *et al.*, *Nature 308*, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 88*, 2825, 1984), and, preferably, differential display (Liang & Pardee, *Science 257*, 967-71, 1992; U.S. Patent 5,262,311).

The differential expression information may itself suggest relevant methods for the treatment of disorders involving the human CRIK-sk. For example, treatment may include a modulation of expression of the differentially expressed genes and/or the gene encoding the human CRIK-sk. The differential expression information may indicate whether the expression or activity of the differentially expressed gene or gene product or the human CRIK-sk gene or gene product are up-regulated or down-regulated.

*Screening Methods*

The invention provides assays for screening test compounds that bind to or modulate the activity of a CRIK-sk polypeptide or a CRIK-sk polynucleotide. A test compound preferably binds to a CRIK-sk polypeptide or polynucleotide. More preferably, a test compound decreases or increases enzymatic activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

*Test Compounds*

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. *See* Lam, *Anticancer Drug Des. 12*, 145, 1997.

Methods for the synthesis of molecular libraries are well known in the art (*see*, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6909, 1993; Erb *et al. Proc. Natl. Acad. Sci. U.S.A. 91*, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem. 37*, 2678, 1994; Cho *et al.*, *Science 261*, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2061; Gallop *et al.*, *J. Med. Chem. 37*, 1233, 1994). Libraries of compounds can be presented in solution

(*see, e.g.*, Houghten, *BioTechniques 13*, 412-421, 1992), or on beads (Lam, *Nature 354*, 82-84, 1991), chips (Fodor, *Nature 364*, 555-556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 89*, 1865-1869, 1992), or phage (Scott & Smith, *Science 249*, 386-390, 1990; Devlin, *Science 249*, 404-406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci. 97*, 6378-6382, 1990; Felici, *J. Mol. Biol. 222*, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

*High Throughput Screening*

Test compounds can be screened for the ability to bind to CRIK-sk polypeptides or polynucleotides or to affect CRIK-sk activity or CRIK-sk gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 50 to 500 µl. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

Alternatively, "free format assays," or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 19,* 1614-18 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in

Philadelphia, Pa. (Nov. 7-10, 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity 2*, 57-63 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

*Binding Assays*

For binding assays, the test compound is preferably a small molecule that binds to and occupies, for example, the active site of the CRIK-sk polypeptide, such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules.

In binding assays, either the test compound or the CRIK-sk polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound that is bound to the CRIK-sk polypeptide can then be accomplished, for example, by direct counting of radioemmission, by scintillation

counting, or by determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a CRIK-sk polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a CRIK-sk polypeptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a CRIK-sk polypeptide (McConnell *et al.*, *Science 257*, 1906-1912, 1992).

Determining the ability of a test compound to bind to a CRIK-sk polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA) (Sjolander & Urbaniczky, *Anal. Chem. 63*, 2338-2345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol. 5*, 699-705, 1995). BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another aspect of the invention, a CRIK-sk polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent 5,283,317; Zervos *et al.*, *Cell 72*, 223-232, 1993; Madura *et al.*, *J. Biol. Chem. 268*, 12046-12054, 1993; Bartel *et al.*, *BioTechniques 14*, 920-924, 1993; Iwabuchi *et al.*, *Oncogene 8*, 1693-1696, 1993; and Brent W094/10300), to identify other proteins which bind to or interact with the CRIK-sk polypeptide and modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide

encoding a CRIK-sk polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in vivo* to form an protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein that interacts with the CRIK-sk polypeptide.

It may be desirable to immobilize either the CRIK-sk polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either the CRIK-sk polypeptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach the enzyme polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a CRIK-sk polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, the CRIK-sk polypeptide is a fusion protein comprising a domain that allows the CRIK-sk polypeptide to be bound to a solid support. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed CRIK-sk polypeptide; the mixture is then incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing proteins or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a CRIK-sk polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated CRIK-sk polypeptides (or polynucleotides) or test compounds can be prepared from biotin-NHS(N-hydroxy-succinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to a CRIK-sk polypeptide, polynucleotide, or a test compound, but which do not interfere with a desired binding site, such as the active site of the CRIK-sk polypeptide, can be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the CRIK-sk polypeptide or test compound, enzyme-linked assays which rely on detecting an activity of the CRIK-sk polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a CRIK-sk polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a CRIK-sk polypeptide or polynucleotide can be used in a cell-based assay system. A CRIK-sk polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Binding of the test compound to a CRIK-sk polypeptide or polynucleotide is determined as described above.

*Enzyme Assays*

Test compounds can be tested for the ability to increase or decrease the activity of a human CRIK-sk polypeptide. CRIK-sk activity can be measured, for example, as described in Di Cunto F. *et al.*, J Biol Chem. 1998 Nov 6;273(45):29706-11.

Enzyme assays can be carried out after contacting either a purified CRIK-sk polypeptide, a cell membrane preparation, or an intact cell with a test compound. A test compound that decreases kinase activity of a CRIK-sk polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for decreasing CRIK-sk activity. A test compound which increases a kinase activity of a human CRIK-sk polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for increasing human CRIK-sk activity.

*Gene Expression*

In another embodiment, test compounds that increase or decrease CRIK-sk gene expression are identified. A CRIK-sk polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the CRIK-sk polynucleotide is determined. The level of expression of appropriate mRNA or polypeptide in the presence of the test compound is compared to the level of expression of mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this

comparison. For example, when expression of mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of the mRNA or polypeptide expression. Alternatively, when expression of the mRNA or polypeptide is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of the mRNA or polypeptide expression.

The level of CRIK-sk mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or polypeptide. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a CRIK-sk polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into a CRIK-sk polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell that expresses a CRIK-sk polynucleotide can be used in a cell-based assay system. The CRIK-sk polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

*Pharmaceutical Compositions*

The invention also provides pharmaceutical compositions that can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a CRIK-sk polypeptide, CRIK-sk polynucleotide, ribozymes or antisense oligonucleotides, antibodies which specifically bind to a CRIK-sk polypeptide, or mimetics, activators, or inhibitors of a CRIK-sk polypeptide

activity. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries that facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Pharmaceutical preparations that can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

*Therapeutic Indications and Methods*

Human CRIK-sk can be regulated to treat obesity and COPD.

*Obesity*. Obesity and overweight are defined as an excess of body fat relative to lean body mass. An increase in caloric intake or a decrease in energy expenditure or both can bring about this imbalance leading to surplus energy being stored as fat. Obesity is associated with important medical morbidities and an increase in mortality. The causes of obesity are poorly understood and may be due to genetic factors, environmental factors or a combination of the two to cause a positive energy balance. In contrast, anorexia and cachexia are characterized by an imbalance in energy intake versus energy expenditure leading to a negative energy balance and weight loss. Agents that either increase energy expenditure and/or decrease energy intake,

absorption or storage would be useful for treating obesity, overweight, and associated comorbidities. Agents that either increase energy intake and/or decrease energy expenditure or increase the amount of lean tissue would be useful for treating cachexia, anorexia and wasting disorders.

The hypothalamus has been considered as the feeding control center. Many neuropeptides, hormones, neurotransmitters, etc. that play important roles in he control of energy homoeostasis have been identified in the hypothalamus. See *J. Lip. Res. 40,* 1735-46, 1999; *Pharm. Rev. 52,* 35-61, 2000. Leptin signaling pathway, MC4, and 5-HT2C systems in the hypothalamus play critical roles in the control of body weight homeostasis. Therefore, a gene selectively expressed in the hypothalamus, such as the human CRIK-sk of the invention, is a potential obesity target.

Thus, this gene, translated proteins and agents which modulate this gene or portions of the gene or its products are useful for treating obesity, overweight, anorexia, cachexia, wasting disorders, appetite suppression, appetite enhancement, increases or decreases in satiety, modulation of body weight, and/or other eating disorders such as bulimia. Also this gene, translated proteins and agents which modulate this gene or portions of the gene or its products are useful for treating obesity/overweight-associated comorbidities including hypertension, type 2 diabetes, coronary artery disease, hyperlipidemia, stroke, gallbladder disease, gout, osteoarthritis, sleep apnea and respiratory problems, some types of cancer including endometrial, breast, prostate, and colon cancer, thrombolic disease, polycystic ovarian syndrome, reduced fertility, complications of pregnancy, menstrual irregularities, hirsutism, stress incontinence, and depression.

<u>*COPD.*</u> Chronic obstructive pulmonary (or airways) disease (COPD) is a condition defined physiologically as airflow obstruction that generally results from a mixture of emphysema and peripheral airway obstruction due to chronic bronchitis (Senior & Shapiro, *Pulmonary Diseases and Disorders,* 3d ed., New York, McGraw-Hill, 1998, pp. 659-681, 1998; Barnes, *Chest 117,* 10S-14S, 2000). Emphysema is characterized by destruction of alveolar walls leading to abnormal enlargement of the air spaces of

the lung. Chronic bronchitis is defined clinically as the presence of chronic productive cough for three months in each of two successive years. In COPD, airflow obstruction is usually progressive and is only partially reversible. By far the most important risk factor for development of COPD is cigarette smoking, although the disease does occur in non-smokers.

Chronic inflammation of the airways is a key pathological feature of COPD (Senior & Shapiro, 1998). The inflammatory cell population comprises increased numbers of macrophages, neutrophils, and $CD8^+$ lymphocytes. Inhaled irritants, such as cigarette smoke, activate macrophages which are resident in the respiratory tract, as well as epithelial cells leading to release of chemokines (*e.g.*, interleukin-8) and other chemotactic factors. These chemotactic factors act to increase the neutrophil/-monocyte trafficking from the blood into the lung tissue and airways. Neutrophils and monocytes recruited into the airways can release a variety of potentially damaging mediators such as proteolytic enzymes and reactive oxygen species. Matrix degradation and emphysema, along with airway wall thickening, surfactant dysfunction, and mucus hypersecretion, all are potential sequelae of this inflammatory response that lead to impaired airflow and gas exchange.

Protein kinases are signal transducing enzymes that phosphorylate proteins, including other kinases, and, along with protein phosphatases, regulate the level and extent of protein phosphorylation and activation. Intracellular signalling pathways have important roles in inflammatory processes. These pathways may be activated by cytokines, oxidant stress and other inflammatory mediators (reviewed in Kyraikis and Avruch, 1996 and 2001). For example, the pro-inflammatory cytokines, tumor necrosis factor α (TNFα) and interleukin-1 activate the protein ser/thr kinases c-Jun-NH2-terminal kinase (JNK) and p38 mitogen-activated protein (MAP) kinase, leading to activation of AP-1 and IKB kinase (IKK), which, in turn, leads to activation of the transcription factor NFKB. Activation of NFKB is required for the transcription of several pro-inflammatory molecules, including interleukin-8 and

ICAM-1. Enzymes of the MAP kinase class may also act to increase cytokine production by stabilization of mRNA (Winzen et al., 1999).

Inhibition of specific protein kinases has been shown to elicit anti-inflammatory effects. For example, the accumulation of polymorphonuclear leukocytes in murine lung following intratracheal administration of bacterial lipopolysaccharide can be blocked by inhibition of p38 MAP kinase (Nick, et al. 2000). As a further example, aerosol delivery to rat lungs of antisense oligodeoxynucleotides to syk kinase mRNA, suppressed nitric oxide and TNFα production from alveolar macrophages stimulated with IgG-anti-IgG complexes (Stenton et al. 2000). Protein kinase subtypes are therefore attractive therapeutic targets for the attenuation of the inflammatory response in COPD. See Kyriakis, J.M. and Avruch J. Sounding the alarm: protein kinase cascades activated by stress and inflammation. J Biol Chem 1996, 271:24313-6; Kyriakis, J.M. and Avruch, J. Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. J. Physiol. Rev. 2001, 81:807-69; Winzen, R., Kracht, M., Ritter, B., Wilhelm, A., Chen C.A., Shyu, A., Müller, M., Gaestel, M., Resch, K., and Holtmann, H. The p38 MAP kinase pathway signals for cytokine-induced mRNA stabilization via MAP kinase-activated protein kinase 2 and an AU-rich region-targeted mechanism. EMBO J. 1999, 18: 4969-4980; Nick, J.A., Young, S.K., Brown, K.K., Avdi, N.J., Arndt, P.G., Suratt, B.T., Janes, M.S., Henson, P.M., Worthen, G.S. Role of p38 mitogen-activated protein kinase in a murine model of pulmonary inflammation. J Immunol. 2000, 164:2151-9; and Stenton, G.R., Kim, M.K., Nohara, O., Chen, C.F., Hirji, N., Wills, F.L., Gilchrist, M., Hwang, P.H., Park, J.G., Finlay, W., Jones, R.L., Befus, A.D., Schreiber, A.D. Aerosolized Syk antisense suppresses Syk expression, mediator release from macrophages, and pulmonary inflammation. J Immunol 2000, 164:3790-7.

This invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a modulating agent, an

antisense nucleic acid molecule, a specific antibody, ribozyme, or a CRIK-sk polypeptide binding molecule) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

A reagent which affects CRIK-sk activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce CRIK-sk activity. The reagent preferably binds to an expression product of a human CRIK-sk gene. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells that have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 μg of DNA per 16 nmole of liposome delivered to about $10^6$ cells, more preferably about 1.0 μg of DNA per 16 nmole of liposome delivered to about $10^6$ cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome

delivered to about $10^6$ cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a particular cell type, such as a cell-specific ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods that are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al. Trends in Biotechnol. 11*, 202-05 (1993); Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem. 263*, 621-24 (1988); Wu *et al.*, *J. Biol. Chem. 269*, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 87*, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem. 266*, 338-42 (1991).

*Determination of a Therapeutically Effective Dose*

The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases CRIK-sk activity relative to the CRIK-sk activity which occurs in the absence of the therapeutically effective dose.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

Therapeutic efficacy and toxicity, *e.g.*, $ED_{50}$ (the dose therapeutically effective in 50% of the population) and $LD_{50}$ (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, $LD_{50}/ED_{50}$.

Pharmaceutical compositions that exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the $ED_{50}$ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors that can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and

frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads, protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Effective *in vivo* dosages of an antibody are in the range of about 5 μg to about 50 μg/kg, about 50 μg to about 5 mg/kg, about 100 μg to about 500 μg/kg of patient body weight, and about 200 to about 250 μg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo* dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg, about 1 μg to about 2 mg, about 5 μg to about 500 μg, and about 20 μg to about 100 μg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides that express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a CRIK-sk gene or the activity of a CRIK-sk polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a CRIK-sk gene or the activity of a CRIK-sk polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to CRIK-sk-specific mRNA, quantitative RT-PCR, immunologic detection of a CRIK-sk polypeptide, or measurement of CRIK-sk activity.

In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described aboye can be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

*Diagnostic Methods*

Human CRIK-sk also can be used in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases and abnormalities related to the presence of mutations in the nucleic acid sequences that encode the enzyme. For example,

differences can be determined between the cDNA or genomic sequence encoding CRIK-sk in individuals afflicted with a disease and in normal individuals. If a mutation is observed in some or all of the afflicted individuals but not in normal individuals, then the mutation is likely to be the causative agent of the disease.

Sequence differences between a reference gene and a gene having mutations can be revealed by the direct DNA sequencing method. In addition, cloned DNA segments can be employed as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. For example, a sequencing primer can be used with a double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures using radiolabeled nucleotides or by automatic sequencing procedures using fluorescent tags.

Genetic testing based on DNA sequence differences can be carried out by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized, for example, by high resolution gel electrophoresis. DNA fragments of different sequences can be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (*see, e.g.*, Myers *et al., Science 230,* 1242, 1985). Sequence changes at specific locations can also be revealed by nuclease protection assays, such as RNase and S 1 protection or the chemical cleavage method (*e.g.*, Cotton *et al., Proc. Natl. Acad. Sci. USA 85,* 4397-4401, 1985). Thus, the detection of a specific DNA sequence can be performed by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes and Southern blotting of genomic DNA. In addition to direct methods such as gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

Altered levels of CRIK-sk also can be detected in various tissues. Assays used to detect levels of the receptor polypeptides in a body sample, such as blood or a tissue biopsy, derived from a host are well known to those of skill in the art and include radioimmunoassays, competitive binding assays, Western blot analysis, and ELISA assays.

All patents and patent applications cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples, which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

**EXAMPLE 1**

*Detection of human citron rho/rac-interacting kinase-short kinase activity*

Subconfluent COS7 cells in 10-cm dishes are transiently transfected by the DEAE-dextran/chloroquine method with 10 µg of FLAG-SEQ ID NO: 1 vector. Cells are harvested 48 h after transfection. Immunoblotting is performed, and cells are probed with anti-FLAG M2 antibodies (Eastman Kodak Co.) Blots are developed using horseradish peroxidase-conjugated secondary antibodies and ECL detection system (Amersham Pharmacia Biotech). In vitro kinase assays are performed by incubating immune complexes in 50 ml of kinase buffer (50 mM HEPES, pH 7.4, 5 mM MgCl2, 3 mM MnCl2, 1mM dithiothreitol), in the presence or absence of 5 mg of histone H1 or myelin basic protein, plus 0.1 mM ATP and 10 mCi of [gamma-32P] ATP (6000 Ci/mM, NEN Life Science Products) for 30 min at 30°C. The products are analyzed by 5% or 12.5% SDS-PAGE followed by autoradiography. For immunoprecipitation of metabolically labeled proteins, primary keratinocytes are incubated with 0.1 mCi/ml [35S]methionine (Expre35S; NEN Life Science Products) for 4 h in methionine-free medium in the presence of serum. Immunoprecipitated proteins are separated on a 5% SDS-PAGE gel and visualized by autoradiography. It is shown

that the polypeptide of SEQ ID NO: 2 has a human citron rho/rac-interacting kinase-short kinase activity.

**EXAMPLE 2**

*Expression of recombinant human CRIK-sk*

The *Pichia pastoris* expression vector pPICZB (Invitrogen, San Diego, CA) is used to produce large quantities of recombinant human CRIK-sk polypeptides in yeast. The CRIK-sk-encoding DNA sequence is derived from SEQ ID NO: 1. Before insertion into vector pPICZB, the DNA sequence is modified by well known methods in such a way that it contains at its 5'-end an initiation codon and at its 3'-end an enterokinase cleavage site, a His6 reporter tag and a termination codon. Moreover, at both termini recognition sequences for restriction endonucleases are added and after digestion of the multiple cloning site of pPICZ B with the corresponding restriction enzymes the modified DNA sequence is ligated into pPICZB. This expression vector is designed for inducible expression in *Pichia pastoris*, driven by a yeast promoter. The resulting pPICZ/md-His6 vector is used to transform the yeast.

The yeast is cultivated under usual conditions in 5 liter shake flasks and the recombinantly produced protein isolated from the culture by affinity chromatography (Ni-NTA-Resin) in the presence of 8 M urea. The bound polypeptide is eluted with buffer, pH 3.5, and neutralized. Separation of the polypeptide from the His6 reporter tag is accomplished by site-specific proteolysis using enterokinase (Invitrogen, San Diego, CA) according to manufacturer's instructions. Purified human CRIK-sk polypeptide is obtained.

**EXAMPLE 3**

*Identification of test compounds that bind to CRIK-sk polypeptides*

Purified CRIK-sk polypeptides comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well microtiter plates are contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. Human CRIK-sk polypeptides comprise the amino acid sequence shown in SEQ ID NO: 2. The test compounds comprise a fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a CRIK-sk polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound that increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound is not incubated is identified as a compound which binds to a CRIK-sk polypeptide.

**EXAMPLE 4**

*Identification of a test compound which decreases CRIK-sk gene expression*

A test compound is administered to a culture of human cells transfected with a CRIK-sk expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells that have not been transfected is incubated for the same time without the test compound to provide a negative control.

RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem. 18*, 5294-99, 1979). Northern blots are prepared using 20 to 30 μg total RNA and hybridized with a $^{32}$P-labeled CRIK-sk-specific probe at 65°C in Express-hyb

(CLONTECH). The probe comprises at least 11 contiguous nucleotides selected from the complement of SEQ ID NO: 1. A test compound that decreases the CRIK-sk-specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of CRIK-sk gene expression.

**EXAMPLE 5**

*Identification of a test compound which decreases CRIK-sk activity*

A test compound is administered to a culture of human cells transfected with a CRIK-sk expression construct and incubated at 37 °C for 10 to 45 minutes. A culture of the same type of cells that have not been transfected is incubated for the same time without the test compound to provide a negative control. CRIK-sk activity is measured using the method of Di Cunto F. *et al.*, J Biol Chem. 1998 Nov 6;273(45):29706-11.

A test compound which decreases the CRIK-sk activity of the CRIK-sk relative to the CRIK-sk activity in the absence of the test compound is identified as an inhibitor of CRIK-sk activity.

**EXAMPLE 6**

*Tissue-specific expression of CRIK-sk*

The qualitative expression pattern of CRIK-sk in various tissues is determined by Reverse Transcription-Polymerase Chain Reaction (RT-PCR).

To demonstrate that CRIK-sk is involved in the disease process of obesity, expression is determined in the following tissues: subcutaneous adipose tissue, mesenteric adipose tissue, adrenal gland, bone marrow, brain (cerebellum, spinal cord, cerebral cortex, caudate, medulla, substantia nigra, and putamen), colon, fetal

brain, heart, kidney, liver, lung, mammary gland, pancreas, placenta, prostate, salivary gland, skeletal muscle small intestine, spleen, stomach, testes, thymus, thyroid trachea, and uterus. Neuroblastoma cell lines SK-Nr-Be (2), Hr, Sk-N-As, HTB-10, IMR-32, SNSY-5Y, T3, SK-N-D2, D283, DAOY, CHP-2, U87MG, BE(2)C, T986, KANTS, MO59K, CHP234, C6 (rat), SK-N-F1, SK-PU-DW, PFSK-1, BE(2)M17, and MCIXC also are tested for CRIK-sk expression. As a final step, the expression of CRIK-sk in cells derived from normal individuals with the expression of cells derived from obese individuals is compared.

To demonstrate that CRIK-sk is involved in the disease process of COPD, the initial expression panel consists of RNA samples from respiratory tissues and inflammatory cells relevant to COPD: lung (adult and fetal), trachea, freshly isolated alveolar type II cells, cultured human bronchial epithelial cells, cultured small airway epithelial cells, cultured bronchial sooth muscle cells, cultured H441 cells (Clara-like), freshly isolated neutrophils and monocytes, and cultured monocytes (macrophage-like). Body map profiling also is carried out, using total RNA panels purchased from Clontech. The tissues are adrenal gland, bone marrow, brain, colon, heart, kidney, liver, lung, mammary gland, pancreas, prostate, salivary gland, skeletal muscle, small intestine, spleen, stomach, testis, thymus, trachea, thyroid, and uterus.

*Quantitative expression profiling.* Quantitative expression profiling is performed by the form of quantitative PCR analysis called "kinetic analysis" firstly described in Higuchi *et al.*, *BioTechnology 10*, 413-17, 1992, and Higuchi *et al.*, *BioTechnology 11*, 1026-30, 1993. The principle is that at any given cycle within the exponential phase of PCR, the amount of product is proportional to the initial number of template copies.

If the amplification is performed in the presence of an internally quenched fluorescent oligonucleotide (TaqMan probe) complementary to the target sequence, the probe is cleaved by the 5'-3' endonuclease activity of Taq DNA polymerase and a fluorescent dye released in the medium (Holland *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*

*88*, 7276-80, 1991). Because the fluorescence emission will increase in direct proportion to the amount of the specific amplified product, the exponential growth phase of PCR product can be detected and used to determine the initial template concentration (Heid *et al.*, *Genome Res. 6*, 986-94, 1996, and Gibson *et al.*, *Genome Res. 6*, 995-1001, 1996).

The amplification of an endogenous control can be performed to standardize the amount of sample RNA added to a reaction. In this kind of experiment, the control of choice is the 18S ribosomal RNA. Because reporter dyes with differing emission spectra are available, the target and the endogenous control can be independently quantified in the same tube if probes labeled with different dyes are used.

All "real time PCR" measurements of fluorescence are made in the ABI Prism 7700.

*RNA extraction and cDNA preparation.* Total RNA from the tissues listed above are used for expression quantification. RNAs labeled "from autopsy" were extracted from autoptic tissues with the TRIzol reagent (Life Technologies, MD) according to the manufacturer's protocol.

Fifty µg of each RNA were treated with DNase I for 1 hour at 37°C in the following reaction mix: 0.2 U/µl RNase-free DNase I (Roche Diagnostics, Germany); 0.4 U/µl RNase inhibitor (PE Applied Biosystems, CA); 10 mM Tris-HCl pH 7.9; 10 mM $MgCl_2$; 50 mM NaCl; and 1 mM DTT.

After incubation, RNA is extracted once with 1 volume of phenol:chloroform:-isoamyl alcohol (24:24:1) and once with chloroform, and precipitated with 1/10 volume of 3 M NaAcetate, pH5.2, and 2 volumes of ethanol.

Fifty µg of each RNA from the autoptic tissues are DNase treated with the DNA-free kit purchased from Ambion (Ambion, TX). After resuspension and spectrophotometric quantification, each sample is reverse transcribed with the TaqMan

Reverse Transcription Reagents (PE Applied Biosystems, CA) according to the manufacturer's protocol. The final concentration of RNA in the reaction mix is 200 ng/μL. Reverse transcription is carried out with 2.5μM of random hexamer primers.

*TaqMan quantitative analysis*. Specific primers and probe are designed according to the recommendations of PE Applied Biosystems; the probe can be labeled at the 5' end FAM (6-carboxy-fluorescein) and at the 3' end with TAMRA (6-carboxy-tetramethyl-rhodamine). Quantification experiments are performed on 10 ng of reverse transcribed RNA from each sample. Each determination is done in triplicate.

Total cDNA content is normalized with the simultaneous quantification (multiplex PCR) of the 18S ribosomal RNA using the Pre-Developed TaqMan Assay Reagents (PDAR) Control Kit (PE Applied Biosystems, CA).

The assay reaction mix is as follows: 1X final TaqMan Universal PCR Master Mix (from 2X stock) (PE Applied Biosystems, CA); 1X PDAR control – 18S RNA (from 20X stock); 300 nM forward primer; 900 nM reverse primer; 200 nM probe; 10 ng cDNA; and water to 25 μl.

Each of the following steps are carried out once: pre PCR, 2 minutes at 50°C, and 10 minutes at 95°C. The following steps are carried out 40 times: denaturation, 15 seconds at 95°C, annealing/extension, 1 minute at 60°C.

The experiment is performed on an ABI Prism 7700 Sequence Detector (PE Applied Biosystems, CA). At the end of the run, fluorescence data acquired during PCR are processed as described in the ABI Prism 7700 user's manual in order to achieve better background subtraction as well as signal linearity with the starting target quantity.

**EXAMPLE 7**

*Identification of test compound efficacy in a COPD animal model*

Guinea pigs are exposed on a single occasion to tobacco smoke for 50 minutes. Animals are sacrificed between 10 minutes and 24 hour following the end of the exposure and their lungs placed in RNAlater™. The lung tissue is homogenised ,and total RNA IS extracted using a Qiagen RNeasy™ Maxi kit. Molecular Probes RiboGreen™ RNA quantitation method is used to quantify the amount of RNA in each sample.

Total RNA is reverse transcribed, and the resultant cDNA is used in a real-time polymerase chain reaction (PCR). The cDNA is added to a solution containing the sense and anti-sense primers and the 6-carboxy-tetramethyl-rhodamine labeled probe of the CRIK-sk gene. Cyclophilin is used as the housekeeping gene. The expression of the CRIK-sk gene is measured using the TaqMan real-time PCR system that generates an amplification curve for each sample. From this curve a threshold cycle value is calculated: the fractional cycle number at which the amount of amplified target reaches a fixed threshold. A sample containing many copies of the CRIK-sk gene will reach this threshold earlier than a sample containing fewer copies. The threshold is set at 0.2, and the threshold cycle $C_T$ is calculated from the amplification curve. The $C_T$ value for the CRIK-sk gene is normalized using the $C_T$ value for the housekeeping gene.

Expression of the CRIK-sk gene is increased by at least 3-fold between 10 minutes and 3 hours post tobacco smoke exposure compared to air exposed control animals.

Test compounds are evaluated as follows. Animals are pre-treated with a test compound between 5 minutes and 1 hour prior to the tobacco smoke exposure and they are then sacrificed up to 3 hours after the tobacco smoke exposure has been completed. Control animals are pre-treated with the vehicle of the test compound via

the route of administration chosen for the test compound. A test compound that reduces the tobacco smoke induced upregulation of CRIK-sk gene relative to the expression seen in vehicle treated tobacco smoke exposed animals is identified as an inhibitor of CRIK-sk gene expression.

**EXAMPLE 8**

*Expression of human citron rho/rac-interacting kinase-short kinase*

Total RNA used for Taqman quantitative analysis were either purchased (Clontech, CA) or extracted from tissues using TRIzol reagent (Life Technologies, MD) according to a modified vendor protocol which utilizes the Rneasy protocol (Qiagen, CA). One hundred µg of each RNA were treated with DNase I using RNase free-DNase (Qiagen, CA) for use with RNeasy or QiaAmp columns.

After elution and quantitation with Ribogreen (Molecular Probes Inc., OR), each sample was reverse transcribed using the GibcoBRL Superscript II First Strand Synthesis System for RT-PCR according to vendor protocol (Life Technologies, MD). The final concentration of RNA in the reaction mix was 50 ng/µL. Reverse transcription was performed with 50 ng of random hexamers.

Specific primers and probe were designed according to PE Applied Biosystems' Primer Express program recommendations and are listed below:

forward primer: 5'-( TCATCAAAAGCAAAGAGCTACAAGA)-3'
reverse primer: 5'-( CATATACGGACGGGAGGATCCT)-3'
probe: SYBR Green

Quantitation experiments were performed on 25 ng of reverse transcribed RNA from each sample. 18S ribosomal RNA was measured as a control using the Pre-

Developed TaqMan Assay Reagents (PDAR)(PE Applied Biosystems, CA). The assay reaction mix was as follows:

| | final concentration/amount |
|---|---|
| TaqMan SYBR Green PCR Master Mix (2x) (PE Applied Biosystems, CA) | 1x |
| Forward primer | 300 nM |
| Reverse primer | 300 nM |
| cDNA | 25 ng |
| Water to 25 uL | |

PCR conditions:

Once: 2' minutes at 50° C

10 minutes at 95°C

40cycles: 15 sec.at 95°C

1 minute at 60°C

The experiment was performed on an ABI Prism 7700 Sequence Detector (PE Applied Biosystems, CA). At the end of the run, fluorescence data acquired during PCR were processed as described in the ABI Prism 7700 user's manual. Fold change was calculated using the delta-delta CT method with normalization to the 18S values. Relative expression was calculated by normalizing to 18s (D Ct), then making the highest expressing tissue 100 and everything else relative to it. Copy number conversion was performed without normalization using the formula Cn=10(Ct-40.007)/-3.623.

The results are shown in FIG. 17.

## REFERENCES

1. Di Cunto F, Imarisio S, Hirsch E, Broccoli V, Bulfone A, Migheli A, Atzori C, Turco E, Triolo R, Dotto GP, Silengo L, Altruda F. Defective neurogenesis in citron kinase knockout mice by altered cytokinesis and massive apoptosis. Neuron. 2000 Oct;28(1):115-27.

2. Di Cunto F, Calautti E, Hsiao J, Ong L, Topley G, Turco E, Dotto GP. Citron rho-interacting kinase, a novel tissue-specific ser/thr kinase encompassing the Rho-Rac-binding protein Citron. J Biol Chem. 1998 Nov 6;273(45):29706-11.

3. Madaule P, Furuyashiki T, Reid T, Ishizaki T, Watanabe G, Morii N, Narumiya S. A novel partner for the GTP-bound forms of rho and rac. FEBS Lett. 1995 Dec 18;377(2):243-8.

4. Fujisawa K, Madaule P, Ishizaki T, Watanabe G, Bito H, Saito Y, Hall A, Narumiya S. Different regions of Rho determine Rho-selective binding of different classes of Rho target molecules. J Biol Chem. 1998 Jul 24;273(30):18943-9.

## CLAIMS

1. An isolated polynucleotide being selected from the group consisting of:

a. a polynucleotide encoding a human citron rho/rac-interacting kinase-short kinase polypeptide comprising an amino acid sequence selected form the group consisting of:

i. amino acid sequences which are at least about 88% identical to the amino acid sequence shown in SEQ ID NO: 2;

ii. the amino acid sequence shown in SEQ ID NO: 2;

iii. amino acid sequences which are at least about 88% identical to the amino acid sequence shown in SEQ ID NO: 9; and

iv. the amino acid sequence shown in SEQ ID NO: 9.

b. a polynucleotide comprising the sequence of SEQ ID NOS: 1 or 8;

c. a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b) and encodes a human citron rho/rac-interacting kinase-short kinase polypeptide;

d. a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code and encodes a human citron rho/rac-interacting kinase-short kinase polypeptide; and

e. a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d) and encodes a human citron rho/rac-interacting kinase-short kinase polypeptide.

2. An expression vector containing any polynucleotide of claim 1.

3. A host cell containing the expression vector of claim 2.

4. A substantially purified human citron rho/rac-interacting kinase-short kinase polypeptide encoded by a polynucleotide of claim 1.

5. A method for producing a human citron rho/rac-interacting kinase-short kinase polypeptide, wherein the method comprises the following steps:

a. culturing the host cell of claim 3 under conditions suitable for the expression of the human citron rho/rac-interacting kinase-short kinase polypeptide; and

b. recovering the human citron rho/rac-interacting kinase-short kinase polypeptide from the host cell culture.

6. A method for detection of a polynucleotide encoding a human citron rho/rac-interacting kinase-short kinase polypeptide in a biological sample comprising the following steps:

a. hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and

b. detecting said hybridization complex.

7. The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.

8. A method for the detection of a polynucleotide of claim 1 or a human citron rho/rac-interacting kinase-short kinase polypeptide of claim 4 comprising the steps of:

a. contacting a biological sample with a reagent which specifically interacts with the polynucleotide or the human citron rho/rac-interacting kinase-short kinase polypeptide and

b. detecting the interaction

9. A diagnostic kit for conducting the method of any one of claims 6 to 8.

10. A method of screening for agents which decrease the activity of a human citron rho/rac-interacting kinase-short kinase, comprising the steps of:

a. contacting a test compound with any human citron rho/rac-interacting kinase-short kinase polypeptide encoded by any polynucleotide of claim1;

b. detecting binding of the test compound to the human citron rho/rac-interacting kinase-short kinase polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential therapeutic agent for decreasing the activity of a human citron rho/rac-interacting kinase-short kinase.

11. A method of screening for agents which regulate the activity of a human citron rho/rac-interacting kinase-short kinase, comprising the steps of:

a. contacting a test compound with a human citron rho/rac-interacting kinase-short kinase polypeptide encoded by any polynucleotide of claim 1; and

b. detecting a human citron rho/rac-interacting kinase-short kinase activity of the polypeptide, wherein a test compound which increases the human citron rho/rac-interacting kinase-short kinase activity is identified as a potential therapeutic agent for increasing the activity of the human citron rho/rac-interacting kinase-short kinase, and wherein a test compound which decreases the human citron rho/rac-interacting kinase-short kinase activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the human citron rho/rac-interacting kinase-short kinase.

12. A method of screening for agents which decrease the activity of a human citron rho/rac-interacting kinase-short kinase, comprising the steps of:

a. contacting a test compound with any polynucleotide of claim 1 and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of human citron rho/rac-interacting kinase-short kinase.

13. A method of reducing the activity of human citron rho/rac-interacting kinase-short kinase, comprising the steps of:

a. contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any human citron rho/rac-interacting kinase-short kinase polypeptide of claim 4, whereby the activity of human citron rho/rac-interacting kinase-short kinase is reduced.

14. A reagent that modulates the activity of a human citron rho/rac-interacting kinase-short kinase polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.

15. A pharmaceutical composition, comprising:

a. the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.

16. Use of the expression vector of claim 2 or the reagent of claim 14 in the preparation of a medicament for modulating the activity of a human citron rho/rac-interacting kinase-short kinase in a disease.

17. Use of claim 16 wherein the disease is obesity or COPD.

Fig. 1

```
atgttgaagt tcaaatatgg agcgcggaat cctttggatg ctggtgctgc tgaacccatt      60
gccagccggg cctccaggct gaatctgttc ttccagggga aaccaccctt tatgactcaa     120
cagcagatgt ctcctctttc ccgagaaggg atattagatg ccctctttgt tctctttgaa     180
gaatgcagtc agcctgctct gatgaagatt aagcacgtga gcaactttgt ccggaagtat     240
tccgacacca tagctgagtt acaggagctc cagccttcgg caaaggactt cgaagtcaga     300
agtcttgtag gttgtggtca ctttgctgaa gtgcaggtgg taagagagaa agcaaccggg     360
gacatctatg ctatgaaagt gatgaagaag aaggctttat tggcccagga gcaggtttca     420
ttttttgagg aagagcggaa catattatct cgaagcacaa gcccgtggat cccccaatta     480
cagtatgcct ttcaggacaa aaatcacctt tatctggtca tggaatatca gcctggaggg     540
gacttgctgt cacttttgaa tagatatgag gaccagttag atgaaaacct gatacagttt     600
tacctagctg agctgatttt ggctgttcac agcgttcatc tgatgggata cgtgcatcga     660
gacatcaagc ctgagaacat tctcgttgac cgcacaggac acatcaagct ggtggatttt     720
ggatctgccg cgaaaatgaa ttcaaacaag atggtgaatg ccaaactccc gattgggacc     780
ccagattaca tggctcctga agtgctgact gtgatgaacg gggatggaaa aggcacctac     840
ggcctggact gtgactggtg gtcagtgggc gtgattgcct atgagatgat ttatgggaga     900
tccccccttcg cagagggaac ctctgccaga accttcaata acattatgaa tttccagcgg     960
tttttgaaat ttccagatga ccccaaagtg agcagtgact ttcttgatct gattcaaagc    1020
ttgttgtgcg gccagaaaga gagactgaag tttgaaggtc tttgctgcca tcctttcttc    1080
tctaaaattg actggaacaa cattcgtaac tctcctcccc ccttcgttcc caccctcaag    1140
tctgacgatg acacctccaa ttttgatgaa ccagagaaga attcgtgggt ttcatcctct    1200
ccgtgccagc tgagcccctc aggcttctcg ggtgaagaac tgccgtttgt ggggttttcg    1260
tacagcaagg cactggggat tcttggtaga tctgagtctg ttgtgtcggg tctggactcc    1320
cctgccaaga ctagctccat ggaaaagaaa cttctcatca aaagcaaaga gctacaagac    1380
tctcaggaca agtgtcacaa ggtatttatt tccgcagccg gcctccttcc ttgctccagg    1440
atcctcccgt ccgtatatgc caagggatcc gcccggggcc gctgc                    1485
```

Fig. 2

```
Met Leu Lys Phe Lys Tyr Gly Ala Arg Asn Pro Leu Asp Ala Gly Ala
 1               5                  10                  15
Ala Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
            20                  25                  30
Gly Lys Pro Pro Phe Met Thr Gln Gln Gln Met Ser Pro Leu Ser Arg
        35                  40                  45
Glu Gly Ile Leu Asp Ala Leu Phe Val Leu Phe Glu Glu Cys Ser Gln
    50                  55                  60
Pro Ala Leu Met Lys Ile Lys His Val Ser Asn Phe Val Arg Lys Tyr
65                  70                  75                  80
Ser Asp Thr Ile Ala Glu Leu Gln Glu Leu Gln Pro Ser Ala Lys Asp
                85                  90                  95
Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
            100                 105                 110
Val Val Arg Glu Lys Ala Thr Gly Asp Ile Tyr Ala Met Lys Val Met
        115                 120                 125
Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
    130                 135                 140
Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
145                 150                 155                 160
Gln Tyr Ala Phe Gln Asp Lys Asn His Leu Tyr Leu Val Met Glu Tyr
                165                 170                 175
Gln Pro Gly Gly Asp Leu Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
            180                 185                 190
```

Fig. 2 (continued)

```
Leu Asp Glu Asn Leu Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
        195                 200                 205
Val His Ser Val His Leu Met Gly Tyr Val His Arg Asp Ile Lys Pro
    210                 215                 220
Glu Asn Ile Leu Val Asp Arg Thr Gly His Ile Lys Leu Val Asp Phe
225                 230                 235                 240
Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Met Val Asn Ala Lys Leu
                245                 250                 255
Pro Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met
            260                 265                 270
Asn Gly Asp Gly Lys Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser
        275                 280                 285
Val Gly Val Ile Ala Tyr Glu Met Ile Tyr Gly Arg Ser Pro Phe Ala
    290                 295                 300
Glu Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg
305                 310                 315                 320
Phe Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Asp Phe Leu Asp
                325                 330                 335
Leu Ile Gln Ser Leu Leu Cys Gly Gln Lys Glu Arg Leu Lys Phe Glu
            340                 345                 350
Gly Leu Cys Cys His Pro Phe Phe Ser Lys Ile Asp Trp Asn Asn Ile
        355                 360                 365
Arg Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp
    370                 375                 380
```

Fig. 2 (continued)

```
Thr Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Val Ser Ser Ser
385                 390                 395                 400
Pro Cys Gln Leu Ser Pro Ser Gly Phe Ser Gly Glu Glu Leu Pro Phe
                405                 410                 415
Val Gly Phe Ser Tyr Ser Lys Ala Leu Gly Ile Leu Gly Arg Ser Glu
            420                 425                 430
Ser Val Val Ser Gly Leu Asp Ser Pro Ala Lys Thr Ser Ser Met Glu
        435                 440                 445
Lys Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys
    450                 455                 460
Cys His Lys Val Phe Ile Ser Ala Ala Gly Leu Leu Pro Cys Ser Arg
465                 470                 475                 480
Ile Leu Pro Ser Val Tyr Ala Lys Gly Ser Ala Arg Gly Arg Cys
                485                 490                 495
```

Fig. 3

```
Met Leu Lys Phe Lys Tyr Gly Val Arg Asn Pro Pro Glu Ala Ser Ala
 1               5                  10                  15
Ser Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
            20                  25                  30
Gly Lys Pro Pro Leu Met Thr Gln Gln Gln Met Ser Ala Leu Ser Arg
        35                  40                  45
Glu Gly Met Leu Asp Ala Leu Phe Ala Leu Phe Glu Glu Cys Ser Gln
    50                  55                  60
Pro Ala Leu Met Lys Met Lys His Val Ser Ser Phe Val Gln Lys Tyr
65                  70                  75                  80
Ser Asp Thr Ile Ala Glu Leu Arg Glu Leu Gln Pro Ser Ala Arg Asp
                85                  90                  95
Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
            100                 105                 110
Val Val Arg Glu Lys Ala Thr Gly Asp Val Tyr Ala Met Lys Ile Met
        115                 120                 125
Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
    130                 135                 140
Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
145                 150                 155                 160
Gln Tyr Ala Phe Gln Asp Lys Asn Asn Leu Tyr Leu Val Met Glu Tyr
                165                 170                 175
Gln Pro Gly Gly Asp Phe Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
            180                 185                 190
```

Fig. 3 (continued)

```
Leu Asp Glu Ser Met Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
        195                 200                 205
Val His Ser Val His Gln Met Gly Tyr Val His Arg Asp Ile Lys Pro
    210                 215                 220
Glu Asn Ile Leu Ile Asp Arg Thr Gly Glu Ile Lys Leu Val Asp Phe
225                 230                 235                 240
Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Val Asp Ala Lys Leu Pro
                245                 250                 255
Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met Asn
            260                 265                 270
Glu Asp Arg Arg Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser Val
        275                 280                 285
Gly Val Val Ala Tyr Glu Met Val Tyr Gly Lys Thr Pro Phe Thr Glu
    290                 295                 300
Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg Phe
305                 310                 315                 320
Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Glu Leu Leu Asp Leu
                325                 330                 335
Leu Gln Ser Leu Leu Cys Val Gln Lys Glu Arg Leu Lys Phe Glu Gly
            340                 345                 350
Leu Cys Cys His Pro Phe Phe Ala Arg Thr Asp Trp Asn Asn Ile Arg
        355                 360                 365
Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp Thr
    370                 375                 380
```

Fig. 3 (continued)

```
Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Ala Phe Ile Leu Cys
385                 390                 395                 400
Val Pro Ala Glu Pro Leu Ala Phe Ser Gly Glu Glu Leu Pro Phe Val
                405                 410                 415
Gly Phe Ser Tyr Ser Lys Ala Leu Gly Tyr Leu Gly Arg Ser Glu Ser
            420                 425                 430
Val Val Ser Ser Leu Asp Ser Pro Ala Lys Val Ser Ser Met Glu Lys
        435                 440                 445
Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys Cys
    450                 455                 460
His Lys Val Ser Ile Ser Thr Ala Gly Leu Arg Pro Cys Ser Arg Ile
465                 470                 475                 480
Leu Gln Ser Ile Tyr Ala Glu Gly Ser Ala Gly Gly His Cys
                485                 490
```

Fig. 4

```
Met Leu Lys Phe Lys Tyr Gly Val Arg Asn Pro Pro Glu Ala Ser Ala
 1               5                  10                  15
Ser Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
            20                  25                  30
Gly Lys Pro Pro Leu Met Thr Gln Gln Gln Met Ser Ala Leu Ser Arg
        35                  40                  45
Glu Gly Met Leu Asp Ala Leu Phe Ala Leu Phe Glu Glu Cys Ser Gln
    50                  55                  60
Pro Ala Leu Met Lys Met Lys His Val Ser Ser Phe Val Gln Lys Tyr
65                  70                  75                  80
Ser Asp Thr Ile Ala Glu Leu Arg Glu Leu Gln Pro Ser Ala Arg Asp
                85                  90                  95
Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
            100                 105                 110
Val Val Arg Glu Lys Ala Thr Gly Asp Val Tyr Ala Met Lys Ile Met
        115                 120                 125
Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
    130                 135                 140
Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
145                 150                 155                 160
Gln Tyr Ala Phe Gln Asp Lys Asn Asn Leu Tyr Leu Val Met Glu Tyr
                165                 170                 175
Gln Pro Gly Gly Asp Phe Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
            180                 185                 190
```

Fig. 4 (continued)

```
Leu Asp Glu Ser Met Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
        195                 200                 205
Val His Ser Val His Gln Met Gly Tyr Val His Arg Asp Ile Lys Pro
    210                 215                 220
Glu Asn Ile Leu Ile Asp Arg Thr Gly Glu Ile Lys Leu Val Asp Phe
225                 230                 235                 240
Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Val Asp Ala Lys Leu Pro
                245                 250                 255
Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met Asn
            260                 265                 270
Glu Asp Arg Arg Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser Val
        275                 280                 285
Gly Val Val Ala Tyr Glu Met Val Tyr Gly Lys Thr Pro Phe Thr Glu
    290                 295                 300
Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg Phe
305                 310                 315                 320
Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Glu Leu Leu Asp Leu
                325                 330                 335
Leu Gln Ser Leu Leu Cys Val Gln Lys Glu Arg Leu Lys Phe Glu Gly
            340                 345                 350
Leu Cys Cys His Pro Phe Phe Ala Arg Thr Asp Trp Asn Asn Ile Arg
        355                 360                 365
Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp Thr
    370                 375                 380
```

Fig. 4 (continued)

```
Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Ala Phe Ile Leu Cys
385                 390                 395                 400
Val Pro Ala Glu Pro Leu Ala Phe Ser Gly Glu Glu Leu Pro Phe Val
                405                 410                 415
Gly Phe Ser Tyr Ser Lys Ala Leu Gly Tyr Leu Gly Arg Ser Glu Ser
            420                 425                 430
Val Val Ser Ser Leu Asp Ser Pro Ala Lys Val Ser Ser Met Glu Lys
        435                 440                 445
Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys Cys
    450                 455                 460
His Lys Met Glu Gln Glu Met Thr Arg Leu His Arg Arg Val Ser Glu
465                 470                 475                 480
Val Glu Ala Val Leu Ser Gln Lys Glu Val Glu Leu Lys Ala Ser Glu
                485                 490                 495
Thr Gln Arg Ser Leu Leu Glu Gln Asp Leu Ala Thr Tyr Ile Thr Glu
            500                 505                 510
Cys Ser Ser Leu Lys Arg Ser Leu Glu Gln Ala Arg Met Glu Val Ser
        515                 520                 525
Gln Glu Asp Asp Lys Ala Leu Gln Leu Leu His Asp Ile Arg Glu Gln
    530                 535                 540
Ser Arg Lys Leu Gln Glu Ile Lys Glu Gln Glu Tyr Gln Ala Gln Val
545                 550                 555                 560
Glu Glu Met Arg Leu Met Met Asn Gln Leu Glu Glu Asp Leu Val Ser
                565                 570                 575
```

Fig. 4 (continued)

```
Ala Arg Arg Arg Ser Asp Leu Tyr Glu Ser Glu Leu Arg Glu Ser Arg
            580                 585                 590
Leu Ala Ala Glu Glu Phe Lys Arg Lys Ala Asn Glu Cys Gln His Lys
        595                 600                 605
Leu Met Lys Ala Lys Asp Gln Gly Lys Pro Glu Val Gly Glu Tyr Ser
    610                 615                 620
Lys Leu Glu Lys Ile Asn Ala Glu Gln Gln Leu Lys Ile Gln Glu Leu
625                 630                 635                 640
Gln Glu Lys Leu Glu Lys Ala Val Lys Ala Ser Thr Glu Ala Thr Glu
                645                 650                 655
Leu Leu Gln Asn Ile Arg Gln Ala Lys Glu Arg Ala Glu Arg Glu Leu
            660                 665                 670
Glu Lys Leu His Asn Arg Glu Asp Ser Ser Glu Gly Ile Lys Lys Lys
        675                 680                 685
Leu Val Glu Ala Glu Glu Arg Arg His Ser Leu Glu Asn Lys Val Lys
    690                 695                 700
Arg Leu Glu Thr Met Glu Arg Arg Glu Asn Arg Leu Lys Asp Asp Ile
705                 710                 715                 720
Gln Thr Lys Ser Glu Gln Ile Gln Gln Met Ala Asp Lys Ile Leu Glu
                725                 730                 735
Leu Glu Glu Lys His Arg Glu Ala Gln Val Ser Ala Gln His Leu Glu
            740                 745                 750
Val His Leu Lys Gln Lys Glu Gln His Tyr Glu Glu Lys Ile Lys Val
        755                 760                 765
```

Fig. 4 (continued)

```
Leu Asp Asn Gln Ile Lys Lys Asp Leu Ala Asp Lys Glu Ser Leu Glu
    770                 775                 780
Asn Met Met Gln Arg His Glu Glu Glu Ala His Glu Lys Gly Lys Ile
785                 790                 795                 800
Leu Ser Glu Gln Lys Ala Met Ile Asn Ala Met Asp Ser Lys Ile Arg
                805                 810                 815
Ser Leu Glu Gln Arg Ile Val Glu Leu Ser Glu Ala Asn Lys Leu Ala
            820                 825                 830
Ala Asn Ser Ser Leu Phe Thr Gln Arg Asn Met Lys Ala Gln Glu Glu
        835                 840                 845
Met Ile Ser Glu Leu Arg Gln Gln Lys Phe Tyr Leu Glu Thr Gln Ala
    850                 855                 860
Gly Lys Leu Glu Ala Gln Asn Arg Lys Leu Glu Glu Gln Leu Glu Lys
865                 870                 875                 880
Ile Ser His Gln Asp His Ser Asp Lys Ser Arg Leu Leu Glu Leu Glu
                885                 890                 895
Thr Arg Leu Arg Glu Val Ser Leu Glu His Glu Glu Gln Lys Leu Glu
            900                 905                 910
Leu Lys Arg Gln Leu Thr Glu Leu Gln Leu Ser Leu Gln Glu Arg Glu
        915                 920                 925
Ser Gln Leu Thr Ala Leu Gln Ala Ala Arg Ala Ala Leu Glu Ser Gln
    930                 935                 940
Leu Arg Gln Ala Lys Thr Glu Leu Glu Glu Thr Thr Ala Glu Ala Glu
945                 950                 955                 960
```

Fig. 4 (continued)

```
Glu Glu Ile Gln Ala Leu Thr Ala His Arg Asp Glu Ile Gln Arg Lys
                965                 970                 975
Phe Asp Ala Leu Arg Asn Ser Cys Thr Val Ile Thr Asp Leu Glu Glu
            980                 985                 990
Gln Leu Asn Gln Leu Thr Glu Asp Asn Ala Glu Leu Asn Asn Gln Asn
        995                 1000                1005
Phe Tyr Leu Ser Lys Gln Leu Asp Glu Ala Ser Gly Ala Asn Asp Glu
    1010                1015                1020
Ile Val Gln Leu Arg Ser Glu Val Asp His Leu Arg Arg Glu Ile Thr
1025                1030                1035                1040
Glu Arg Glu Met Gln Leu Thr Ser Gln Lys Gln Thr Met Glu Ala Leu
                1045                1050                1055
Lys Thr Thr Cys Thr Met Leu Glu Glu Gln Val Leu Asp Leu Glu Ala
            1060                1065                1070
Leu Asn Asp Glu Leu Leu Glu Lys Glu Arg Gln Trp Glu Ala Trp Arg
        1075                1080                1085
Ser Val Leu Gly Asp Glu Lys Ser Gln Phe Glu Cys Arg Val Arg Glu
    1090                1095                1100
Leu Gln Arg Met Leu Asp Thr Glu Lys Gln Ser Arg Ala Arg Ala Asp
1105                1110                1115                1120
Gln Arg Ile Thr Glu Ser Arg Gln Val Val Glu Leu Ala Val Lys Glu
                1125                1130                1135
His Lys Ala Glu Ile Leu Ala Leu Gln Gln Ala Leu Lys Glu Gln Lys
            1140                1145                1150
```

Fig. 4 (continued)

```
Leu Lys Ala Glu Ser Leu Ser Asp Lys Leu Asn Asp Leu Glu Lys Lys
        1155                1160                1165
His Ala Met Leu Glu Met Asn Ala Arg Ser Leu Gln Gln Lys Leu Glu
    1170                1175                1180
Thr Glu Arg Glu Leu Lys Gln Arg Leu Leu Glu Glu Gln Ala Lys Leu
1185                1190                1195                1200
Gln Gln Gln Met Asp Leu Gln Lys Asn His Ile Phe Arg Leu Thr Gln
                1205                1210                1215
Gly Leu Gln Glu Ala Leu Asp Arg Ala Asp Leu Leu Lys Thr Glu Arg
            1220                1225                1230
Ser Asp Leu Glu Tyr Gln Leu Glu Asn Ile Gln Val Leu Tyr Ser His
        1235                1240                1245
Glu Lys Val Lys Met Glu Gly Thr Ile Ser Gln Gln Thr Lys Leu Ile
    1250                1255                1260
Asp Phe Leu Gln Ala Lys Met Asp Gln Pro Ala Lys Lys Lys Lys Val
1265                1270                1275                1280
Pro Leu Gln Tyr Asn Glu Leu Lys Leu Ala Leu Glu Lys Glu Lys Ala
                1285                1290                1295
Arg Cys Ala Glu Leu Glu Glu Ala Leu Gln Lys Thr Arg Ile Glu Leu
            1300                1305                1310
Arg Ser Ala Arg Glu Glu Ala Ala His Arg Lys Ala Thr Asp His Pro
        1315                1320                1325
His Pro Ser Thr Pro Ala Thr Ala Arg Gln Gln Ile Ala Met Ser Ala
    1330                1335                1340
```

Fig. 4 (continued)

```
Ile Val Arg Ser Pro Glu His Gln Pro Ser Ala Met Ser Leu Leu Ala
1345                1350                1355                1360
Pro Pro Ser Ser Arg Arg Lys Glu Ser Ser Thr Pro Glu Glu Phe Ser
                1365                1370                1375
Arg Arg Leu Lys Glu Arg Met His His Asn Ile Pro His Arg Phe Asn
            1380                1385                1390
Val Gly Leu Asn Met Arg Ala Thr Lys Cys Ala Val Cys Leu Asp Thr
        1395                1400                1405
Val His Phe Gly Arg Gln Ala Ser Lys Cys Leu Glu Cys Gln Val Met
    1410                1415                1420
Cys His Pro Lys Cys Ser Thr Cys Leu Pro Ala Thr Cys Gly Leu Pro
1425                1430                1435                1440
Ala Glu Tyr Ala Thr His Phe Thr Glu Ala Phe Cys Arg Asp Lys Met
                1445                1450                1455
Asn Ser Pro Gly Leu Gln Ser Lys Glu Pro Gly Ser Ser Leu His Leu
            1460                1465                1470
Glu Gly Trp Met Lys Val Pro Arg Asn Asn Lys Arg Gly Gln Gln Gly
        1475                1480                1485
Trp Asp Arg Lys Tyr Ile Val Leu Glu Gly Ser Lys Val Leu Ile Tyr
    1490                1495                1500
Asp Asn Glu Ala Arg Glu Ala Gly Gln Arg Pro Val Glu Glu Phe Glu
1505                1510                1515                1520
Leu Cys Leu Pro Asp Gly Asp Val Ser Ile His Gly Ala Val Gly Ala
                1525                1530                1535
```

Fig. 4 (continued)

```
Ser Glu Leu Ala Asn Thr Ala Lys Ala Asp Val Pro Tyr Ile Leu Lys
            1540                1545                1550
Met Glu Ser His Pro His Thr Thr Cys Trp Pro Gly Arg Thr Leu Tyr
        1555                1560                1565
Leu Leu Ala Pro Ser Phe Pro Asp Lys Gln Arg Trp Val Thr Ala Leu
    1570                1575                1580
Glu Ser Val Val Ala Gly Gly Arg Val Ser Arg Glu Lys Ala Glu Ala
1585                1590                1595                1600
Asp Ala Lys Leu Leu Gly Asn Ser Leu Leu Lys Leu Glu Gly Asp Asp
                1605                1610                1615
Arg Leu Asp Met Asn Cys Thr Leu Pro Phe Ser Asp Gln Val Val Leu
            1620                1625                1630
Val Gly Thr Glu Glu Gly Leu Tyr Ala Leu Asn Val Leu Lys Asn Ser
        1635                1640                1645
Leu Thr His Ile Pro Gly Ile Gly Ala Val Phe Gln Ile Tyr Ile Ile
    1650                1655                1660
Lys Asp Leu Glu Lys Leu Leu Met Ile Ala Gly Glu Glu Arg Ala Leu
1665                1670                1675                1680
Cys Leu Val Asp Val Lys Lys Val Lys Gln Ser Leu Ala Gln Ser His
                1685                1690                1695
Leu Pro Ala Gln Pro Asp Val Ser Pro Asn Ile Phe Glu Ala Val Lys
            1700                1705                1710
Gly Cys His Leu Phe Ala Ala Gly Lys Ile Glu Asn Ser Leu Cys Ile
        1715                1720                1725
```

Fig. 4 (continued)

```
Cys Ala Ala Met Pro Ser Lys Val Val Ile Leu Arg Tyr Asn Asp Asn
    1730                1735                1740
Leu Ser Lys Tyr Cys Ile Arg Lys Glu Ile Glu Thr Ser Glu Pro Cys
1745                1750                1755                1760
Ser Cys Ile His Phe Thr Asn Tyr Ser Ile Leu Ile Gly Thr Asn Lys
                1765                1770                1775
Phe Tyr Glu Ile Asp Met Lys Gln Tyr Thr Leu Asp Glu Phe Leu Asp
            1780                1785                1790
Lys Asn Asp His Ser Leu Ala Pro Ala Val Phe Ala Ser Ser Ser Asn
        1795                1800                1805
Ser Phe Pro Val Ser Ile Val Gln Ala Asn Ser Ala Gly Gln Arg Glu
    1810                1815                1820
Glu Tyr Leu Leu Cys Phe His Glu Phe Gly Val Phe Val Asp Ser Tyr
1825                1830                1835                1840
Gly Arg Arg Ser Arg Thr Asp Asp Leu Lys Trp Ser Arg Leu Pro Leu
                1845                1850                1855
Ala Phe Ala Tyr Arg Glu Pro Tyr Leu Phe Val Thr His Phe Asn Ser
            1860                1865                1870
Leu Glu Val Ile Glu Ile Gln Ala Arg Ser Ser Leu Gly Ser Pro Ala
        1875                1880                1885
Arg Ala Tyr Leu Glu Ile Pro Asn Pro Arg Tyr Leu Gly Pro Ala Ile
    1890                1895                1900
Ser Ser Gly Ala Ile Tyr Leu Ala Ser Ser Tyr Gln Asp Lys Leu Arg
1905                1910                1915                1920
```

Fig. 4 (continued)

```
Val Ile Cys Cys Lys Gly Asn Leu Val Lys Glu Ser Gly Thr Glu Gln
                1925                1930                1935
His Arg Val Pro Ser Thr Ser Arg Ser Ser Pro Asn Lys Arg Gly Pro
            1940                1945                1950
Pro Thr Tyr Asn Glu His Ile Thr Lys Arg Val Ala Ser Ser Pro Ala
        1955                1960                1965
Pro Pro Glu Gly Pro Ser His Pro Arg Glu Pro Ser Thr Pro His Arg
    1970                1975                1980
Tyr Arg Asp Arg Glu Gly Arg Thr Glu Leu Arg Arg Asp Lys Ser Pro
1985                1990                1995                2000
Gly Arg Pro Leu Glu Arg Glu Lys Ser Pro Gly Arg Met Leu Ser Thr
                2005                2010                2015
Arg Arg Glu Arg Ser Pro Gly Arg Leu Phe Glu Asp Ser Ser Arg Gly
            2020                2025                2030
Arg Leu Pro Ala Gly Ala Val Arg Thr Pro Leu Ser Gln Val Asn Lys
        2035                2040                2045
Val Trp Asp Gln Ser Ser Val
    2050                2055
```

Fig. 5

```
gggcggggct gagggcggcg ggggcgggcc gcccgagctg ggagggcggc ggcgccgagg      60
ggaggagagc ggcccatgga cccgcggggc ccggcgcccc agactctgcg ccgtcgggac     120
ggagcccaag atgtcggcct aggccggggc gcgacgacgc ggacggggcg gcgaggaggc     180
gccgctgctg ccggggctcg cagccgccga gcccccgagg gcgcgccctg acggactggc     240
cgagccggcg gtgagaggcc ggcgcgtctg gagcgggccg cgcggcacca tgtcggccaa     300
ggtgcggctc aagaagctgg agcagctgct cctggacggg ccctggcgca acgagagcgc     360
cctgagcgtg gaaacgctgc tcgacgtgct cgtctgcctg tacaccgagt gcagccactc     420
ggccctgcgc cgcgacaagt acgtggccga gttcctcgag tgggctaaac catttacaca     480
gctggtgaaa gaaatgcagc ttcatcgaga agactttgaa ataattaaag taattggaag     540
aggtgctttt ggtgaggttg ctgttgtcaa aatgaagaat actgaacgaa tttatgcaat     600
gaaaatcctc aacaagtggg agatgctgaa aagagcagag accgcgtgct tccgagagga     660
gcgcgatgtg ctggtgaacg gcgactgcca gtggatcacc gcgctgcact acgcctttca     720
ggacgagaac cacctgtact tagtcatgga ttactatgtg ggtggtgatt tactgaccct     780
gctcagcaaa tttgaagaca agcttccgga agatatggcg aggttctaca ttggtgaaat     840
ggtgctggcc attgactcca tccatcagct tcattacgtg cacagagaca ttaaacctga     900
caatgtcctt ttggacgtga atggtcatat ccgcctggct gactttggat catgtttgaa     960
gatgaatgat gatggcactg tgcagtcctc cgtggccgtg ggcacacctg actacatctc    1020
gccggagatc ctgcaggcga tggaggacgg catgggcaaa tacgggcctg agtgtgactg    1080
gtggtctctg ggtgtctgca tgtatgagat gctctatgga gaaacgccgt tttatgcgga    1140
gtcactcgtg gagacctatg ggaagatcat gaaccatgaa gagcgattcc agttcccatc    1200
ccatgtcacg gatgtatctg aagaagcgaa ggacctcatc cagagactga tctgcagtag    1260
agaacgccgg ctggggcaga atggaataga ggatttcaaa aagcatgcgt tttttgaagg    1320
tctaaattgg gaaaatatac gaaacctaga agcaccttat attcctgatg tgagcagtcc    1380
```

Fig. 5 (continued)

```
ctctgacaca tccaacttcg acgtggatga cgacgtgctg agaaacacgg aaatattacc    1440
tcctggttct cacacaggct tttctggatt acatttgcca ttcattggtt ttacattcac    1500
aacggaaagc tgtttttctg atcgaggctc tctgaagagc ataatgcagt ccaacacatt    1560
aaccaaagat gaggatgtgc agcgggacct ggagcacagc ctgcagatgg aagcttacga    1620
gaggaggatt cggaggctgg aacaggagaa gctggagctg agcaggaagc tgcaagagtc    1680
cacccagacc gtgcagtccc tccacggctc atctcgggcc ctcagcaatt caaaccgaga    1740
taaagaaatc aaaaagctaa atgaagaaat cgaacgcttg aagaataaaa tagcagattc    1800
aaacaggctg gagcgacagc ttgaggacac agtggcgctt cgccaagagc gtgaggactc    1860
cacgcagcgg ctgcgggggc tggagaagca gcaccgcgtg gtccggcagg agaaggagga    1920
gctgcacaag caactggttg aagcctcaga gcggttgaaa tcccaggcca aggaactcaa    1980
agatgcccat cagcagcgaa agctggccct gcaggagttc tcggagctga acgagcgcat    2040
ggcagagctc cgtgcccaga agcagaaggt gtcccggcag ctgcgagaca aggaggagga    2100
gatggaggtg gccacgcaga aggtggacgc catgcggcag gaaatgcgga gagctgagaa    2160
gctcaggaaa gagctggaag ctcagcttga tgatgctgtt gctgaggcct ccaaggagcg    2220
caagcttcgt gagcacagcg agaacttctg caagcaaatg gaaagcgagc tggaggccct    2280
caaggtgaag caaggaggcc ggggagcggg tgccacctta gagcaccagc aagagatttc    2340
caaaatcaaa tccgagctgg agaagaaagt cttattttat gaagaggaat tggtcagacg    2400
tgaggcctcc catgtgctag aagtgaaaaa tgtgaagaag gaggtgcatg attcagaaag    2460
ccaccagctg gccctgcaga aagaaatctt gatgttaaaa gataagttag aaaagtcaaa    2520
gcgagaacgg cataacgaga tggaggaggc agtaggtaca ataaaagata aatacgaacg    2580
agaaagagcg atgctgtttg atgaaaacaa gaagctaact gctgaaaatg aaaagctctg    2640
ttcctttgtg gataaactca cagctcaaaa tagacagctg gaggatgagc tgcaggatct    2700
ggcagccaag aaggagtcag tggcccactg ggaagctcag attgcggaaa tcattcagtg    2760
```

Fig. 5 (continued)

```
ggtcagtgac gagaaagatg cccggggtta ccttcaagct cttgcttcca agatgaccga    2820
agagctcgag gctttgagga gttctagtct ggggtcaaga acactggacc cgctgtggaa    2880
ggtgcgccgc agccagaagc tggacatgtc cgcgcggctg gagctgcagt cggccctgga    2940
ggcggagatc cgggccaagc agcttgtcca ggaggagctc aggaaggtca aggacgccaa    3000
cctcaccttg gaaagcaaac taaaggattc cgaagccaaa aacagagaat tattagaaga    3060
aatggaaatt ttgaagaaaa agatggaaga aaaattcaga gcagatactg ggctcaaact    3120
tccagatttt caggattcca tttttgagta tttcaacact gctcctcttg cacatgacct    3180
gacatttaga accagctcag ctagtgagca agaaacacaa gctccgaagc cagaagcgtc    3240
cccgtcgatg tctgtggctg catcagagca gcaggaggac atggctcggc ccccgcagag    3300
gccatccgct gtgccgttgc ccaccacgca ggccctggtt ctggctggac cgaagccaaa    3360
agctcaccag ttcagcatca agtccttctc cagccctact cagtgcagcc actgcacctc    3420
cctgatggtt gggctgatcc ggcagggcta cgcctgcgag gtgtgttcct ttgcttgcca    3480
cgtgtcctgc aaagacggtg ccccccaggt gtgcccaata cctcccgagc agtccaagag    3540
gcctctgggc gtggacgtgc agcgaggcat cggaacagcc tacaaaggcc atgtcaaggt    3600
cccaaagccc acgggggtga agaagggatg gcagcgcgca tatgcagtcg tctgtgagtg    3660
caagctcttc ctgtatgatc tgcctgaagg aaaatccacc cagcctggtg tcattgcgag    3720
ccaagtcttg gatctcagag atgacgagtt ttccgtgagc tcagtcctgg cctcagatgt    3780
cattcatgct acacgccgag atattccatg tatattcagg gtgacggcct ctctcttagg    3840
tgcaccttct aagaccagct cgctgctcat tctgacagaa aatgagaatg aaaagaggaa    3900
gtgggttggg attctagaag gactccagtc catccttcat aaaaaccggc tgaggaatca    3960
ggtcgtgcat gttcccttgg aagcctacga cagctcgctg cctctcatca aggccatcct    4020
gacagctgcc atcgtggatg cagacaggat tgcagtcggc ctagaagaag ggctctatgt    4080
catagaggtc acccgagatg tgatcgtccg tgccgctgac tgtaagaagg tacaccagat    4140
```

Fig. 5 (continued)

```
cgagcttgct cccagggaga agatcgtaat cctcctctgt ggccggaacc accatgtgca    4200
cctctatccg tggtcgtccc ttgatggagc ggaaggcagc tttgacatca agcttccgga    4260
aaccaaaggc tgccagctca tggccacggc cacactcaag aggaactctg gcacctgcct    4320
gtttgtggcc gtgaaacggc tgatcctttg ctatgagatc cagagaacga agccattcca    4380
cagaaagttc aatgagattg tggctcccgg cagcgtgcag tgcctggcgg tgctcaggga    4440
caggctctgt gtgggctacc cttctgggtt ctgcctgctg agcatccagg gggacgggca    4500
gcctctaaac ctggtaaatc ccaatgaccc ctcgcttgcg ttcctctcac aacagtcttt    4560
tgatgccctt tgtgctgtgg agctcgaaag cgaggagtac ctgctttgct tcagccacat    4620
gggactgtac gtggacccgc aaggccggag ggcacgcgcg caggagctca tgtggcctgc    4680
ggctcctgtc gcctgtagtt gcagccccac ccacgtcacg gtgtacagcg agtatggcgt    4740
ggacgtcttt gatgtgcgca ccatggagtg ggtgcagacc atcggcctgc ggaggataag    4800
gcccctgaac tctgaaggca ccctcaacct cctcaactgc gagcctccac gcttgatcta    4860
cttcaagagc aagttctcgg gagcggttct caacgtgccg gacacctccg acaacagcaa    4920
gaagcagatg ctgcgcacca ggagcaaaag gcggttcgtc ttcaaggtcc cagaggaaga    4980
gagactgcag cagaggcgag agatgcttag agacccagaa ttgagatcca aaatgatatc    5040
caacccaacc aacttcaacc acgtggccca catgggccca ggcgacggca tgcaggtgct    5100
catggacctg cctctgagtg ctgtgccccc ctcccaggag gaaaggccgg gccccgctcc    5160
caccaacctg gctcgccagc ctccatccag gaacaagccc tacatctcgt ggccctcatc    5220
aggtggatcg gagcctagcg tgactgtgcc tctgagaagt atgtctgatc cagaccagga    5280
ctttgacaaa gagcctgatt cggactccac caaacactca actccatcga atagctccaa    5340
ccccagcggc ccaccgagcc ccaactcccc ccacaggagc cagctccccc tcgaaggcct    5400
ggagcagccg gcctgtgaca cctgaagccg ccagctcgcc acaggggcca gggagctgga    5460
gatggcctcc agcgtcagtg ccaagactga gcgggccctc cagtgttgtc caaggaaatg    5520
```

Fig. 5 (continued)

```
tagaatcact ttgtagatat ggagatgaag aagacaaatc tttattataa tattgatcag    5580
ttttatgccg cattgttcgt ggcagtagac cacatctgtt cgtctgcaca gctgtgaggc    5640
gatgctgttc catctgcaca tgaaggaccc ccatacagcc tgtctcccac ccctgacaac    5700
ccgagagggc atatggggcc ctgccaacac cacttcctca gcagaaaccc gtcatgacgc    5760
ggctgcttcg gaagcagaca tctggggaca cagcctcagt acccagtctt ttccctagtt    5820
cctgaaactt tcctaggacc ttaagagaat agtaggaggt cctatagcat tcccagtgtc    5880
actagaattt tgaagacagg aaagtggagg ttagtctgtg gccttttttt catttagcca    5940
ttgcacagtc agctgcagaa gtcctgctga ccacctagtc atggacaaag gcccaggacc    6000
agtgacaccc tgcgtccctg tgtgcgttaa gttcattctg ggtcgcagcc atgaagtgtc    6060
accagtatct actactgtga agtcagctgt gctgttttcc attcgcttcc acggcttctg    6120
cctcctgcca taaaaccagc gagtgtcgtg gtgcaggcag gccctgtggc ctgctgggct    6180
gagggaagtc agagccccag ggcgccacga agcagccact gggataccc accccgcccc    6240
gccnncccccc ccccccccnc cagtcnagnc ccgaaatgga gcccccgtga ttagtagccc    6300
gtatgatcac gtagacccac ccaacacact cctgcacact ggccccggcc cacggcacag    6360
caatcccctg cgcgtggatt tcacctcacc ctttgtacca gatgttgagt gaccagctct    6420
gtggccctgt gtcgtcagag gcttgtgatt aactgtggcg gcagacacag cttgtccaca    6480
gcttgggcca ggcttccct gtcctcccac cggtcggctg cttggcaagg ctgttcagga    6540
cgtgcacttc cccaagtcgg cactgagtgg cccagcacca cctagccctg ccacccccact    6600
gccctcctgg gccttctgct ggatgggcac ctgggggtt ctggtttttt actttttaa    6660
tgtaagtctc agtctttgta attaattatt gaattgtgag aacatttttg aacaatttac    6720
ctgtcaataa agcagaagac ggcagtttta aagttaaaaa aaaaaaaaaa aaaaaaaaaa    6780
```

Fig. 6

```
gagcggccgc ccgggcaggt ctagaattca gcggccgctg aattctaggt gctgccggag      60
acctcagggc cccttaaaga ggacccattc ccctgtagac cagtctctgt cccctgcaag     120
ctttacctgc attcttgccc atggcgcttc ccattttctg tgtgatctta tctccctcac     180
ttaactctct gttcctgtgt cttcattcta tgagctggac tgaggccttg ggtgggaaag     240
tgggctctgt attctattcc gtgcctaacc cagcgcctcc ttcttgtgtc tttctccctc     300
tctagcctat ctggtcagtc aggcaaccga tcttcctcag gatcattgat ctctgtacct     360
ccaggggcag tgaaccttcc tttccctggg ataatcctca aggctcactg atcaaacctt     420
tgggcttggt tcacaggtta ggtctatgtc agtacgcgac atcagatatt tgtgttcgtc     480
agggttttct aggggaaaag agctggtaga atggaaaagt ggagatttat taggctgcag     540
tctgctagtc caccaatggc tggtagttct tttggaaatg atttatttcc atcccttatg     600
tgtatgagta cctttggcct tcctgtgtct gtgccccatg tgccgtggag cgtggtcgcc     660
acctcctcat cctgatctct ttagggagac acgactctgc caagcccttc ctgccttcaa     720
tgtcagtacc cgcttgactt tccccagtgt tccttcggcg ttggcggaga gatgttgaag     780
ttcaagtatg gtgtgcggaa cccgccggag gccagtgcct ccgagcccat tgccagtcgg     840
gcctccaggc taaatctctt cttccagggg aaaccgcccc tcatgactca acagcagatg     900
tctgctcttt cccgggaagg gatgctagac gccctcttcg ctctctttga agagtgcagc     960
caacccgccc tgatgaagat gaagcacgtg agcagctttg tccagaagta ttccgacacc    1020
atagccgagt tgcgggagct gcagccgtcg gcgagagact tcgaagttcg aagccttgtg    1080
ggctgtggtc acttcgctga agtgcaggtg gttagagaga aggcgaccgg ggacgtctat    1140
gccatgaaaa tcatgaagaa gaaggctttg ctggcccagg aacaggtttc atttttcgag    1200
gaggagagga acatattatc tcggagcacg agtccttgga tcccccagtt acagtacgcc    1260
tttcaggaca aaaataacct ttacctggtc atggaatatc agcctggagg ggatttcctg    1320
tcgcttctga acagatacga ggaccaatta gatgagagca tgatccagtt ttaccttgct    1380
```

Fig. 6 (continued)

```
gagctgattt tggctgtcca cagcgtgcac cagatgggat atgtgcatcg agacatcaag    1440
cccgagaaca tcctcatcga ccggacggga gagatcaagc tggtggattt tggatcagcc    1500
gctaagatga attcaaataa ggtggatgcc aaactcccca ttgggacccc ggattacatg    1560
gctccggaag tgttgaccgt gatgaacgag gaccgaaggg gcacatacgg cttggactgt    1620
gactggtggt ctgtcggagt tgttgcttat gagatggttt atgggaagac cccattcaca    1680
gagggaacct ccgcccggac cttcaacaac atcatgaact tccagcggtt tttgaagttc    1740
ccagatgacc ccaaagttag cagtgagctc cttgatctgc ttcagagtct gctgtgtgtc    1800
cagaaagaga gactgaagtt cgagggtctc tgctgccacc ctttctttgc cagaacggac    1860
tggaacaaca tccgtaactc tcctcccccc ttcgtcccca ccctcaagtc tgacgatgac    1920
acctccaatt ttgatgaacc agagaagaat tcgtgggctt tcatcctctg tgtgccagct    1980
gagcccctcg cgttctcagg cgaagagctg ccgtttgtgg gattttcgta cagcaaggca    2040
ctggggtatc ttggtagatc tgagtctgtc gtgtcgagtc tggactcccc tgccaaggtt    2100
agctccatgg aaaagaaact tctcatcaaa agcaaagagc tccaagactc ccaggacaag    2160
tgtcacaagg tatctatctc cacagccggc ctccgtcctt gctccaggat cctccagtca    2220
atatatgccg agggatctgc cgggggccac tgctgagcgg tggtgccgcc tccctcgctg    2280
aagtcgtgcc tccagcagct cagagggaga ggactccagg cccagacatt gccataaatc    2340
ctttaaatct taaccagagg aggccctgga tttaaaaaaa                          2380
```

Fig. 7

```
Met Ser Ala Glu Val Arg Leu Arg Gln Leu Gln Gln Leu Val Leu Asp
 1               5                  10                  15
Pro Gly Phe Leu Gly Leu Glu Pro Leu Leu Asp Leu Leu Leu Gly Val
            20                  25                  30
His Gln Glu Leu Gly Ala Ser His Leu Ala Gln Asp Lys Tyr Val Ala
        35                  40                  45
Asp Phe Leu Gln Trp Val Glu Pro Ile Ala Ala Arg Leu Lys Glu Val
    50                  55                  60
Arg Leu Gln Arg Asp Asp Phe Glu Ile Leu Lys Val Ile Gly Arg Gly
65                  70                  75                  80
Ala Phe Ser Glu Val Ala Val Val Lys Met Lys Gln Thr Gly Gln Val
                85                  90                  95
Tyr Ala Met Lys Ile Met Asn Lys Trp Asp Met Leu Lys Arg Gly Glu
            100                 105                 110
Val Ser Cys Phe Arg Glu Glu Arg Asp Val Leu Val Lys Gly Asp Arg
        115                 120                 125
Arg Trp Ile Thr Gln Leu His Phe Ala Phe Gln Asp Glu Asn Tyr Leu
    130                 135                 140
Tyr Leu Val Met Glu Tyr Tyr Val Gly Gly Asp Leu Leu Thr Leu Leu
145                 150                 155                 160
Ser Lys Phe Gly Glu Arg Ile Pro Ala Glu Met Ala Arg Phe Tyr Leu
                165                 170                 175
Ala Glu Ile Val Met Ala Ile Asp Ser Val His Arg Leu Gly Tyr Val
            180                 185                 190
```

Fig. 7 (continued)

```
His Arg Asp Ile Lys Pro Asp Asn Ile Leu Leu Asp Arg Cys Gly His
        195                 200                 205
Ile Arg Leu Ala Asp Phe Gly Ser Cys Leu Lys Leu Gln Pro Asp Gly
    210                 215                 220
Met Val Arg Ser Leu Val Ala Val Gly Thr Pro Asp Tyr Leu Ser Pro
225                 230                 235                 240
Glu Ile Leu Gln Ala Val Gly Gly Gly Pro Gly Ala Gly Ser Tyr Gly
                245                 250                 255
Pro Glu Cys Asp Trp Trp Ala Leu Gly Val Phe Ala Tyr Glu Met Phe
            260                 265                 270
Tyr Gly Gln Thr Pro Phe Tyr Ala Asp Ser Thr Ala Glu Thr Tyr Ala
        275                 280                 285
Lys Ile Val His Tyr Arg Glu His Leu Ser Leu Pro Leu Ala Asp Thr
    290                 295                 300
Val Val Pro Glu Glu Ala Gln Asp Leu Ile Arg Gly Leu Leu Cys Pro
305                 310                 315                 320
Ala Glu Ile Arg Leu Gly Arg Gly Gly Ala Gly Asp Phe Gln Lys His
                325                 330                 335
Pro Phe Phe Phe Gly Leu Asp Trp Glu Gly Leu Arg Asp Ser Val Pro
            340                 345                 350
Pro Phe Thr Pro Asp Phe Glu Gly Ala Thr Asp Thr Cys Asn Phe Asp
        355                 360                 365
Val Val Glu Asp Arg Leu Thr Ala Met Val Ser Gly Gly Gly Glu Thr
    370                 375                 380
```

Fig. 7 (continued)

```
Leu Ser Asp Met Gln Glu Asp Met Pro Leu Gly Val Arg Leu Pro Phe
385                 390                 395                 400
Val Gly Tyr Ser Tyr Cys Cys Met Ala Phe Arg Asp Asn Gln Val Pro
                405                 410                 415
Asp Pro Thr Pro Met Glu Leu Glu Ala Leu Gln Leu Pro Val Ser Asp
            420                 425                 430
Leu Gln Gly Leu Asp Leu Gln Pro Pro Val Ser Pro Pro Asp Gln Val
        435                 440                 445
Ala Glu Glu Ala Asp Leu Val Ala Val Pro Ala Pro Val Ala Glu Ala
    450                 455                 460
Glu Thr Thr Val Thr Leu Gln Gln Leu Gln Glu Ala Leu Glu Glu Glu
465                 470                 475                 480
Val Leu Thr Arg Gln Ser Leu Ser Arg Glu Leu Glu Ala Ile Arg Thr
                485                 490                 495
Ala Asn Gln Asn Phe Ser Ser Gln Leu Gln Glu Ala Glu Val Arg Asn
            500                 505                 510
Arg Asp Leu Glu Ala His Val Arg Gln Leu Gln Glu Arg Met Glu Met
        515                 520                 525
Leu Gln Ala Pro Gly Ala Ala Ala Ile Thr Gly Val Pro Ser Pro Arg
    530                 535                 540
Ala Thr Asp Pro Pro Ser His Leu Asp Gly Pro Pro Ala Val Ala Val
545                 550                 555                 560
Gly Gln Cys Pro Leu Val Gly Pro Gly Pro Met His Arg Arg His Leu
                565                 570                 575
```

Fig. 7 (continued)

```
Leu Leu Pro Ala Arg Ile Pro Arg Pro Gly Leu Ser Glu Ala Arg Cys
            580                 585                 590
Leu Leu Leu Phe Ala Ala Leu Ala Ala Ala Thr Leu Gly Cys
        595                 600                 605
Thr Gly Leu Val Ala Tyr Thr Gly Gly Leu Thr Pro Val Trp Cys Phe
    610                 615                 620
Pro Gly Ala Thr Phe Ala Pro
625                 630
```

Fig. 8

**BLASTP - alignment of CRIK-sk (SEQ ID NO:2) against trembl|AF086823|AF086823_1** gene: "Crik-sk"; product: "rho/rac-interacting citron kinase short isoform" ; Mus musculus rho/rac-interacting citron kinase short isoform (Crik-sk) mRNA, complete cds. //:gp|AF086823|3599507 gene: "Crik-sk"; product: "rho/rac-interacting citron kinase short isoform" ; Mus musculus rho/rac-interacting citron kinase short isoform (Crik-sk) mRNA, complete cds. (SEQ ID NO:3)

This hit is scoring at : 0.0 (expectation value)
Alignment length (overlap) : 495
Identities : 87 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb_1_;

```
Q:      1 MLKFKYGARNPLDAGAAEPIASRASRLNLFFQGKPPFMTQQQMSPLSREGILDALFVLFE
          MLKFKYG.RNP :A.A:EPIASRASRLNLFFQGKPP.MTQQQMS.LSREG:LDALF.LFE
H:      1 MLKFKYGVRNPPEASASEPIASRASRLNLFFQGKPPLMTQQQMSALSREGMLDALFALFE

                                    Protein_Kinase_ATP Motif (K binds ATP)
          ECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGCGHFAEVQVVREKATG
          ECSQPALMK:KHVS:FV:KYSDTIAEL:ELQPSA:DFEVRSLVGCGHFAEVQVVREKATG
          ECSQPALMKMKHVSSFVQKYSDTIAELRELQPSARDFEVRSLVGCGHFAEVQVVREKATG

          DIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLYLVMEYQPGG
          D:YAMK:MKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKN:LYLVMEYQPGG
          DVYAMKIMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNNLYLVMEYQPGG
```

Fig. 8 (continued)

**Protein_Kinase_ST Motif (D is an active site)**

```
DLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVDF
D.LSLLNRYEDQLDE::IQFYLAELILAVHSVH MGYVHRDIKPENIL:DRTG.IKLVDF
DFLSLLNRYEDQLDESMIQFYLAELILAVHSVHQMGYVHRDIKPENILIDRTGEIKLVDF

GSAAKMNSNKMVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYGR
GSAAKMNSNK V:AKLPIGTPDYMAPEVLTVMN D :GTYGLDCDWWSVGV:AYEM:YG:
GSAAKMNSNK-VDAKLPIGTPDYMAPEVLTVMNEDRRGTYGLDCDWWSVGVVAYEMVYGK

SPFAEGTSARTFNNIMNFQRFLKFPDDPKVSSDFLDLIQSLLCGQKERLKFEGLCCHPFF
:PF.EGTSARTFNNIMNFQRFLKFPDDPKVSS:.LDL:QSLLC QKERLKFEGLCCHPFF
TPFTEGTSARTFNNIMNFQRFLKFPDDPKVSSELLDLLQSLLCVQKERLKFEGLCCHPFF

SKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSSSPCQLSPSGFSGEELPFVGFS
::.DWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSW.    ....P .FSGEELPFVGFS
ARTDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWAFILCVPAEPLAFSGEELPFVGFS

YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKVFISAAGLLPCSR
YSKALG.LGRSESVVS.LDSPAK.SSMEKKLLIKSKELQDSQDKCHKV IS.AGL PCSR
YSKALGYLGRSESVVSSLDSPAKVSSMEKKLLIKSKELQDSQDKCHKVSISTAGLRPCSR

ILPSVYAKGSARGRC     495
IL.S:YA:GSA G.C
ILQSIYAEGSAGGHC     494
```

Fig. 9

BLASTP - alignment of CRIK-sk (SEQ ID NO:2) against trembl|AF086824|AF086824_1
gene: "Crik"; product: "rho/rac-interacting citron kinase"; Mus musculus rho/rac-interacting citron kinase (Crik) mRNA, complete cds.
//:gp|AF086824|3599509 gene: "Crik"; product: "rho/rac-interacting citron kinase"; Mus musculus rho/rac-interacting citron kinase (Crik) mRNA, complete cds. (SEQ ID NO:4)

This hit is scoring at : 0.0 (expectation value)
Alignment length (overlap) : 468
Identities : 88 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb_1_;

```
Q:      1 MLKFKYGARNPLDAGAAEPIASRASRLNLFFQGKPPFMTQQQMSPLSREGILDALFVLFE
          MLKFKYG.RNP :A.A:EPIASRASRLNLFFQGKPP.MTQQQMS.LSREG:LDALF.LFE
H:      1 MLKFKYGVRNPPEASASEPIASRASRLNLFFQGKPPLMTQQQMSALSREGMLDALFALFE

          ECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGCGHFAEVQVVREKATG
          ECSQPALMK:KHVS:FV:KYSDTIAEL:ELQPSA:DFEVRSLVGCGHFAEVQVVREKATG
          ECSQPALMKMKHVSSFVQKYSDTIAELRELQPSARDFEVRSLVGCGHFAEVQVVREKATG

          DIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLYLVMEYQPGG
          D:YAMK:MKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKN:LYLVMEYQPGG
          DVYAMKIMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNNLYLVMEYQPGG
```

Fig. 9 (continued)

```
DLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVDF
D.LSLLNRYEDQLDE::IQFYLAELILAVHSVH MGYVHRDIKPENIL:DRTG.IKLVDF
DFLSLLNRYEDQLDESMIQFYLAELILAVHSVHQMGYVHRDIKPENILIDRTGEIKLVDF

GSAAKMNSNKMVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYGR
GSAAKMNSNK V:AKLPIGTPDYMAPEVLTVMN D :GTYGLDCDWWSVGV:AYEM:YG:
GSAAKMNSNK-VDAKLPIGTPDYMAPEVLTVMNEDRRGTYGLDCDWWSVGVVAYEMVYGK

SPFAEGTSARTFNNIMNFQRFLKFPDDPKVSSDFLDLIQSLLCGQKERLKFEGLCCHPFF
:PF.EGTSARTFNNIMNFQRFLKFPDDPKVSS:.LDL:QSLLC QKERLKFEGLCCHPFF
TPFTEGTSARTFNNIMNFQRFLKFPDDPKVSSELLDLLQSLLCVQKERLKFEGLCCHPFF

SKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSSSPCQLSPSGFSGEELPFVGFS
::.DWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSW.     ....P .FSGEELPFVGFS
ARTDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWAFILCVPAEPLAFSGEELPFVGFS

YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKV      468
YSKALG.LGRSESVVS.LDSPAK.SSMEKKLLIKSKELQDSQDKCHK:
YSKALGYLGRSESVVSSLDSPAKVSSMEKKLLIKSKELQDSQDKCHKM      467
```

Fig. 10

**BLASTP - alignment of CRIK-sk (SEQ ID NO:2) against trembl|AF128625|AF128625_1** gene: "CDC42BPB"; product: "CDC42-binding protein kinase beta"; Homo sapiens CDC42-binding protein kinase beta (CDC42BPB) mRNA, complete cds.
//:gp|AF128625|5006445 gene: "CDC42BPB"; product: "CDC42-binding protein kinase beta"; Homo sapiens CDC42-binding protein kinase beta (CDC42BPB) mRNA, complete cds. (SEQ ID NO:5)

This hit is scoring at : 4e-94 (expectation value)
Alignment length (overlap) : 420
Identities : 42 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb_1_;

```
Q:      44 SPLSREGILDALFVLFEECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLV
           S.LS E :LD.L..L:.ECS..AL.: K:V:.F:.  .. .. ::E:Q ..:DFE: .::
H:      23 SALSVETLLDVLVCLYTECSHSALRRDKYVAEFLEWAKPFTQLVKEMQLHREDFEIIKVI

           GCGHFAEVQVVREKATGDIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYA
           G G F.EV.VV: K T  IYAMK::.K .:L.:.:.: F.EER::L .....WI..L.YA
           GRGAFGEVAVVKMKNTERIYAMKILNKWEMLKRAETACFREERDVLVNGDCQWITALHYA

           FQDKNHLYLVMEYQPGGDLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIK
           FQD:NHLYLVM:Y. GGDLL:LL:::ED:L.E:.:FY:.E::LA:.S:H : YVHRDIK
           FQDENHLYLVMDYYVGGDLLTLLSKFEDKLPEDMARFYIGEMVLAIDSIHQLHYVHRDIK
```

Fig. 10 (continued)

```
PENILVDRTGHIKLVDFGSAAKMNSNKMVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLD
P:N:L:D .GHI:L.DFGS..KMN.: .V.:.:.:GTPDY::PE:L..M. DG.G.YG :
PDNVLLDVNGHIRLADFGSCLKMNDDGTVQSSVAVGTPDYISPEILQAME-DGMGKYGPE

CDWWSVGVIAYEMIYGRSPFAEGTSARTFNNIMNFQRFLKFPDD-PKVSSDFLDLIQSLL
CDWWS:GV..YEM:YG.:PF . : ..T:..IMN.:. .:FP.. ..VS.: DLIQ.L:
CDWWSLGVCMYEMLYGETPFYAESLVETYGKIMNHEERFQFPSHVTDVSEEAKDLIQRLI

CGQKERLKFEGL---CCHPFFSKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSS
C.:.:.RL .G: H.FF. :.W.NIRN ..P::P.:.S..DTSNFD .:. :.
CSRERRLGQNGIEDFKKHAFFEGLNWENIRNLEAPYIPDVSSPSDTSNFDVDDDVLRNTE

SPCQLSPSGFSGEELPFVGFSYSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQ   459
   . S :GFSG .LPF:GF:::        :ES..S  D. : .S M:...L.K.:::Q
ILPPGSHTGFSGLHLPFIGFTFT--------TESCFS--DRGSLKSIMQSNTLTKDEDVQ   431
```

Fig. 11

*BLASTP - alignment of CRIK-sk (SEQ ID NO:2) against swissnew|P54265|DMK_MOUSE*

MYOTONIN-PROTEIN KINASE (EC 2.7.1.-) (MYOTONIC DYSTROPHY PROTEIN KINASE) (MDPK) (DM-KINASE) (DMK) (DMPK) (MT-PK).//:swiss|P54265|DMK_MOUSE MYOTONIN-PROTEIN KINASE (EC 2.7.1.-) (MYOTONIC DISTROPHY PROTEIN KINASE) (MDPK) (DM-KINASE) (DMK) (DMPK) (MT-PK).//:trembl|Z38015|MMMDMPK_1 gene: "DM-PK"; product: "myotonic dystrophy protein kinase"; M.musculus DMR-N9 gene, exons 4 and 5, and DM-PK gene encoding myotonic dystrophy protein kinase //:gp|Z38015|563526 gene: "DM-PK"; product: "myotonic dystrophy protein kinase"; M.musculus DMR-N9 gene, exons 4 and 5, and DM-PK gene encoding myotonic dystrophy protein kinase. (SEQ ID NO:11)

This hit is scoring at : 3e-89 (expectation value)
Alignment length (overlap) : 386
Identities : 44 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb_1_;

```
Q:      46 LSREGILDALFVLFEECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGC
           L. E :LD.L. :.:E.... L.: K:V::F:: .....A.L:E:: ...DFE: .::G
H:      20 LGLEPLLDLLLGVHQELGASHLAQDKYVADFLQWVEPIAARLKEVRLQRDDFEILKVIGR

           GHFAEVQVVREKATGDIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQ
           G F:EV.VV: K.TG.:YAMK:M.K  :L.: :VS F.EER::L :... WI.QL.:AFQ
           GAFSEVAVVKMKQTGQVYAMKIMNKWDMLKRGEVSCFREERDVLVKGDRRWITQLHFAFQ
```

Fig.11 (continued)

```
DKNHLYLVMEYQPGGDLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPE
D:N:LYLVMEY. GGDLL:LL::: :::.....:.:FYLAE:::A:.SVH :GYVHRDIKP:
DENYLYLVMEYYVGGDLLTLLSKFGERIPAEMARFYLAEIVMAIDSVHRLGYVHRDIKPD

NILVDRTGHIKLVDFGSAAKMNSNKMVNAKLPIGTPDYMAPEVL-TVMNGDGKGTYGLDC
NIL:DR.GHI:L.DFGS..K:..: MV.: :.:GTPDY::PE:L .V .G.G.G:YG :C
NILLDRCGHIRLADFGSCLKLQPDGMVRSLVAVGTPDYLSPEILQAVGGGPGAGSYGPEC

DWWSVGVIAYEMIYGRSPFAEGTSARTFNNIMNFQRFLKFP-DDPKVSSDFLDLIQSLLC
DWW::GV.AYEM.YG::PF ..::A.T: .I:::::..L..P  D. V..:  DLI:.LLC
DWWALGVFAYEMFYGQTPFYADSTAETYAKIVHYREHLSLPLADTVVPEEAQDLIRGLLC

GQKERLKFEG---LCCHPFFSKIDWNNIRNSPPPFVPTLKSDDDTSNFD--EPEKNSWVS
 .: RL   G    .  HPFF  :DW..:R:S PPF.P..:. .DT.NFD  E.. .:.VS
PAEIRLGRGGAGDFQKHPFFFGLDWEGLRDSVPPFTPDFEGATDTCNFDVVEDRLTAMVS

SSPCQLSPSGFS---GEELPFVGFSY     421
..  .LS..  .   G .LPFVG:SY
GGGETLSDMQEDMPLGVRLPFVGYSY     405
```

Fig. 12

**BLASTP - alignment of CRIK-sk (SEQ ID NO:2) against pdb|1CDK|1CDK-A**
camp-dependent protein kinase(protein kinase a)protein kinase inhibitor(pki(5-24))

This hit is scoring at : 9e-45 (expectation value)
Alignment length (overlap) : 333
Identities : 33 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb_1_;

```
Q:      71 KHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGCGHFAEVQVVREKATGDIYAMKVMKK
           K  .:F::K:.:.......L.     .FE  ..:G.G.F..V.:V:.K.TG: :AMK::.K
H:      14 KAKEDFLKKWENPAQNTAHLD----QFERIKTLGTGSFGRVMLVKHKETGNHFAMKILDK

           KALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLYLVMEYQPGGDLLSLLNRYE
           :.:::. :Q:..  .E:.IL.. . P:: :L:Y:F:D.::LY:VMEY PGG::.S L.R.
           QKVVKLKQIEHTLNEKRILQAVNFPFLVKLEYSFKDNSNLYMVMEYVPGGEMFSHLRRI-

           DQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVDFGSAAKMNSNK
           .:..E  .:FY.A:::L... :H :..::RD:KPEN:L:D:.G:I:::DFG A.:::...
           GRFSEPHARFYAAQIVLTFEYLHSLDLIYRDLKPENLLIDQQGYIQVTDFGFAKRVKGRT

           MVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYGRSPFAEGTSAR
           .. .    GTP:Y:APE::.      .KG Y.  .DWW::GV:.YEM. G .PF .....:
           WTLC----GTPEYLAPEIIL-----SKG-YNKAVDWWALGVLIYEMAAGYPPFFADQPIQ
```

Fig. 12 (continued)

```
TFNNIMNFQRFLKFPDDPKVSSDFLDLIQSLLCGQKERLKFEG--------LCCHPFFSK
.:..I:: :  ::FP.  ..SSD. DL:::LL  Q : .K  G        :  H.:F:.
IYEKIVSGK--VRFPS--HFSSDLKDLLRNLL--QVDLTKRFGNLKDGVNDIKNHKWFAT

IDWNNI--RNSPPPFVPTLKSDDDTSNFDEPEK     393
.DW  I  R. ..PF:P..K...DTSNFD: E:
TDWIAIYQRKVEAPFIPKFKGPGDTSNFDDYEE     325
```

Fig. 13

**HMMPFAM - alignment of CRIK-sk (SEQ ID NO:2) against pfam|hmm|pkinase**
Protein kinase domain

This hit is scoring at : 219.4 E=5.5e-62
Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:      97 FEVRSLVGCGHFAEVQVVREKaTGDIYAMKVMKKKALLaqeqvsffEEERNILSRSTSPW
           :E:. .:G G.F.:V. .:.K TG.I.A:K::KK::L          .E .IL.R :.P
H:       1 yellek1GeGsfGkVykakhk.tgkivAvKilkkesls.....lrEiqilkrlsHpN

          IPQLQYAFQ-DKNHLYLVMEYQPGGDLLSLLNRYEdQLDENLIQFYLAELILAVHSVHLM
          I :L  .F: ..:HLYLVMEY..GGDL...L.R   .L.E. .: ...::: .:. :H
          Ivrllgvfedtddhlylvm Eymeg GdLfdylrrng.plsekeakkialQilrGleYLHsn

          GYVHRDIKPENILVDRTGHIKLVDFGSAAKMnsnkmVNAKLPIGTPDYM-APEVLtvMNG
          G.VHRD:KPENIL:D..G :K:.DFG A. :      .... :GTP YM APEV:  :.G
          givHRDLKpeNILldengtvKiaDFGLArll.....eklttfvGTpwYmmAPEvi..leg

          dgkGTYGLDCDWWSVGVIAYEMIYG-------------------------RSPFAE---
               Y. ..D WS:GVI.YE:: G                          : PF::
          ...rgysskvDvWSlGviLyEl1tggplfpgadlpaftggdevdqliifvlklPfsdelp
```

Fig. 13 (continued)

```
----GTSARTFNNIMNfqrflKFPDDPKVSSDFLDLIQSLLC-GQKERL---KFEGLCCH
   .. ...F.  ..     :.P ....S.:. DL:...L  ...:R    . : :. H
ktridpleelfrikkr.....rlplpsncSeelkdLlkkcLnkDPskRpGsatakeilnh

PFF     360
P:F
pwf     278
```

Fig. 14

**HMMPFAM - alignment of CRIK-sk (SEQ ID NO:2) against pfam|hmm|pkinase_C**
Protein kinase C terminal domain

This hit is scoring at : 15.4  E=0.0018
Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:    361 SKIDWNNI--RNSPPPFVPTLKSDDDTSNFDE     390
          .:IDW:.:  :. .PPF P.:KS. DTSNFD:
H:      1 reIdWdkLEnkeiePPFKPkiksprDtsNFDk      32
```

Fig. 15

**Prosite search results.**

| | | | |
|---|---|---|---|
| PS00107 | 103->127 | PROTEIN_KINASE_ATP | PDOC00100 |
| PS00108 | 217->230 | PROTEIN_KINASE_ST | PDOC00100 |

Fig. 16

**genewise output:**

```
AF086823_1              1 MLKFKYGVRNPPEASASEPIASRASRLNLFFQ
                          MLKFKYG RNP +A A+EPIASRASRLNLFFQ
                          MLKFKYGARNPLDAGAAEPIASRASRLNLFFQ
gi|13653116|r1909637      atatatggcactgggggggcagacgtacacttc
                          ttataagcgactacgccactcggccgtattta
                          gggcataggttgtttttactccgccggtgccg

AF086823_1             33                                GKPPLMTQQQMSALSREGMLDALFAL
                                                         GKPP MTQQQMS LSREG+LDALF L
                                                         GKPPFMTQQQMSPLSREGILDALFVL
gi|13653116|r1909733      GTAACAG  Intron 1   TAGgacctaacccatcctcggatggctgc
                          0-----[1909733:1916-0>gaccttcaaatcctcgagttactttt
                                                         gaactgtagggtttcaagaatccttc

AF086823_1             59 FEECSQPALMKMKHVSSFVQK                              SD
                          FEECSQPALMK+KHVS+FV+K                              SD
                          FEECSQPALMKIKHVSNFVRK          Y:Y[tat]            SD
gi|13653116|r1916682      tggtaccgcaaaacgaatgcaTGTAAGTT  Intron 2   CAGATtg
                          taaggaccttataatgattga 1-----[1916746:1928-1>  ca
                          taactgttgggtgcgcctcgg                         cc
```

Fig. 16 (continued)

```
AF086823_1              83 TIAELRELQPSARDFEVRSLVGCGHFAEVQVVREKATGDVYAMKIMKKK
                           TIAEL+ELQPSA+DFEVRSLVGCGHFAEVQVVREKATGD+YAMK+MKKK
                           TIAELQELQPSAKDFEVRSLVGCGHFAEVQVVREKATGDIYAMKVMKKK
gi|13653116|r1928115    aaggtcgccctgagtggaacggtgctgggcggagagaggatgaagaaaa
                        ctcataatacccaatatggttgggatcatattgaaccgatactattaaa
                        catgaggcgtgagccacattatttcttagggaagaacgccttgaggggg

AF086823_1             132 ALLAQEQ                          VSFFEEERNILSRSTSPWI
                           ALLAQEQ                          VSFFEEERNILSRSTSPWI
                           ALLAQEQ                          VSFFEEERNILSRSTSPWI
gi|13653116|r1928262    gttgcgcGTAGGAG  Intron 3   TAGgtttgggcaattcaaacta
                        cttcaaa0-----[1928283:1935-0>tcttaaagattcggcgcgt
                        tagcggg                          tattgaggcaatacacggc

AF086823_1             158 PQLQYAFQDKNNLYL                          VMEYQPGGDFL
                           PQLQYAFQDKN+LYL                          VMEYQPGGD+L
                           PQLQYAFQDKNHLYL                          VMEYQPGGDLL
gi|13653116|r1935587    cctctgtcgaacctcGTGAGTC  Intron 4   CAGgagtccgggtc
```

Fig. 16 (continued)

```
                        cataactaaaaatat0-----[1935632:1951-0>ttaaacggatt
                        caagtctgcatcttg                      cgatgtagcgg

AF086823_1         184  SLLNRYEDQLDESMIQFYLAELILAVHSVHQMGYVH
                        SLLNRYEDQLDE++IQFYLAELILAVHSVH MGYVH
                        SLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVH
gi|13653116|r1951610 tctaatggctggacacttcggcatggcagccagtgc
                        cttagaaaataaattatatcatttctagtattgata
                        atgtatgcgatacgagtcatggtgttccttggacgt

AF086823_1         220                          DIKPENILIDRTGEIKLVDFGSA
                                                DIKPENIL+DRTG IKLVDFGSA
                                 R:R[cga]       DIKPENILVDRTGHIKLVDFGSA
gi|13653116|r1951718 CGGTAAGTG  Intron 5   CAGAgaacgaacggcagcaacggtgtg
                          2-----[1951720:1952-2> atacaatttagcgatattatgcc
                                                ccgtgctctccaaccgggttatc
AF086823_1         244  AKMNSNKV                     -DAKLPIGTPDYMAPEVL
                        AKMNSNK+                      AKLPIGTPDYMAPEVL
                        AKMNSNKM                    VNAKLPIGTPDYMAPEVL
gi|13653116|r1953011 gaaataaaGTAAAAA  Intron 6   TAGgagaccagacgtagcggc
                        catacaat0-----[1953035:1960-0>tacatctgccaatccatt
                        gagtacgg                     gtcacgtgcatcgttagg
```

Fig. 16 (continued)

```
AF086823_1          269 TVMNEDRRGTYGLDCDWWSVGVVAYEMVYGKTPFTEGTSARTFNNIMNF
                        TVMN D +GTYGLDCDWWSVGV+AYEM+YG++PF EGTSARTFNNIMNF
                        TVMNGDGKGTYGLDCDWWSVGVIAYEMIYGRSPFAEGTSARTFNNIMNF
gi|13653116|r1960491 agaagggagatgcgtgtttgggagtgaatgatctgggatgaataaaaat
                        cttagagagcagtagaggctgttcaattaggcctcagcccgctaattat
                        tggcgtaaccccgctcggagcgtctggttgacccagactcacctctgtc

AF086823_1          318 Q                         RFLKFPDDPKVSSELLDLLQSLLCV
                        Q                         RFLKFPDDPKVSS+ LDL+QSLLC
                        Q                         RFLKFPDDPKVSSDFLDLIQSLLCG
gi|13653116|r1960638 cGTAAAGA  Intron 7   CAGcttatcggcagaagtcgcacatttg
                        a0-----[1960641:1962-0>gttatcaacatggattattagttgg
                        g                         gtgatatccagctctttgtacggcc

AF086823_1          344 QKERLKFEGLCCHPFFARTDWNNIRN
                        QKERLKFEGLCCHPFF++ DWNNIRN
                        QKERLKFEGLCCHPFFSKIDWNNIRN          S:S[tct]
gi|13653116|r1962909 cagacatggcttcctttaagtaaacaTGTAAGTA  Intron 8
                        aaagtatagtggacttcatagaatga 1-----[1962988:19824
                        gagaggtattccttcctatcgccttc
```

Fig. 16 (continued)

```
AF086823_1          370      PPPFVPTLKSDDDTSNFDEPEKNSWAFILCVPAEPLAFSGEELP
                             PPPFVPTLKSDDDTSNFDEPEKNSW         P  FSGEELP
                             PPPFVPTLKSDDDTSNFDEPEKNSWVSSSPCQLSPSGFSGEELP
gi|13653116|r1982415 CAGCTccctgcacatgggatatggcgaattgtttctccactgttgggcc
                     -1>  cccttcctacaaaccataacaaacgtccccgatgccgtcgaatc
                          tccctcccgtctcccttaaggtggtactgcggccaccgtaagg

AF086823_1          415 FVGFSYSKALGYLGRS                          SVVSSLD
                        FVGFSYSKALG LGRS                          SVVS LD
                        FVGFSYSKALGILGRS            E:E[gag]      SVVSGLD
gi|13653116|r1982552 tggtttaagcgacgatGAGTAAGTG  Intron 9   TAGGtggtgcg
                     ttgtcagactgttggc  2-----[1982602:2000-2> cttcgta
                     tggtgccgaggtttat                          ttggtgc

AF086823_1          439 SPAKVSSMEKKLLIKSKELQDSQDKCHKVSISTAGLRPCSRILQSIYAE
                        SPAK SSMEKKLLIKSKELQDSQDKCHKV IS AGL PCSRIL S+YA+
                        SPAKTSSMEKKLLIKSKELQDSQDKCHKVFISAAGLLPCSRILPSVYAK
gi|13653116|r2000764 tcgaaatagaaccaaaagccgtcgatcagtatgggcccttaacctgtga
                     cccacgctaaatttagaataacaaagaatttcccgttcgcgttcctaca
                     ctcgtccgagatccacagaactgcgtcgattcacccttccgccgcatcg
```

Fig. 16 (continued)

```
AF086823_1          488 GSAGGHC
                        GSA G C
                        GSARGRC
gi|13653116|r2000911    gtgcgct
                        gccgggg
                        accgccc


//
>gi|13653116|ref|NT_009775.3|Hs12_9932.[1909637:2000931].sp
ATGTTGAAGTTCAAATATGGAGCGCGGAATCCTTTGGATGCTGGTGCTGCTGAACCCATT
GCCAGCCGGGCCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCACCCTTTATGACTCAA
CAGCAGATGTCTCCTCTTTCCCGAGAAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA
GAATGCAGTCAGCCTGCTCTGATGAAGATTAAGCACGTGAGCAACTTTGTCCGGAAGTAT
TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAAGGACTTCGAAGTCAGA
AGTCTTGTAGGTTGTGGTCACTTTGCTGAAGTGCAGGTGGTAAGAGAGAAAGCAACCGGG
GACATCTATGCTATGAAAGTGATGAAGAAGAAGGCTTTATTGGCCCAGGAGCAGGTTTCA
TTTTTTGAGGAAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA
CAGTATGCCTTTCAGGACAAAAATCACCTTTATCTGGTCATGGAATATCAGCCTGGAGGG
GACTTGCTGTCACTTTTGAATAGATATGAGGACCAGTTAGATGAAAACCTGATACAGTTT
```

Fig. 16 (continued)

TACCTAGCTGAGCTGATTTTGGCTGTTCACAGCGTTCATCTGATGGGATACGTGCATCGA
GACATCAAGCCTGAGAACATTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT
GGATCTGCCGCGAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC
CCAGATTACATGGCTCCTGAAGTGCTGACTGTGATGAACGGGGATGGAAAAGGCACCTAC
GGCCTGGACTGTGACTGGTGGTCAGTGGGCGTGATTGCCTATGAGATGATTTATGGGAGA
TCCCCCTTCGCAGAGGGAACCTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG
TTTTTGAAATTTCCAGATGACCCCAAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC
TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC
TCTAAAATTGACTGGAACAACATTCGTAACTCTCCTCCCCCCTTCGTTCCCACCCTCAAG
TCTGACGATGACACCTCCAATTTTGATGAACCAGAGAAGAATTCGTGGGTTTCATCCTCT
CCGTGCCAGCTGAGCCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG
TACAGCAAGGCACTGGGGATTCTTGGTAGATCTGAGTCTGTTGTGTCGGGTCTGGACTCC
CCTGCCAAGACTAGCTCCATGGAAAAGAAACTTCTCATCAAAAGCAAAGAGCTACAAGAC
TCTCAGGACAAGTGTCACAAGGTATTTATTTCCGCAGCCGGCCTCCTTCCTTGCTCCAGG
ATCCTCCCGTCCGTATATGCCAAGGGATCCGCCCGGGGCCGCTGC
//

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| gi\|13653116\|ref\|NT_009775.3\|Hs12_9932 | GeneWise | match | 1909637 | 2000931 | 899.93 | + | . |
| AF086823_1 | | | | | | | |
| gi\|13653116\|ref\|NT_009775.3\|Hs12_9932 | GeneWise | cds | 1909637 | 1909732 | 0.00 | + | 0 |
| AF086823_1 | | | | | | | |
| gi\|13653116\|ref\|NT_009775.3\|Hs12_9932 | GeneWise | intron | 1909733 | 1916603 | 0.00 | + | . |
| AF086823_1 | | | | | | | |
| gi\|13653116\|ref\|NT_009775.3\|Hs12_9932 | GeneWise | cds | 1916604 | 1916745 | 0.00 | + | 0 |
| AF086823_1 | | | | | | | |

Fig. 16 (continued)

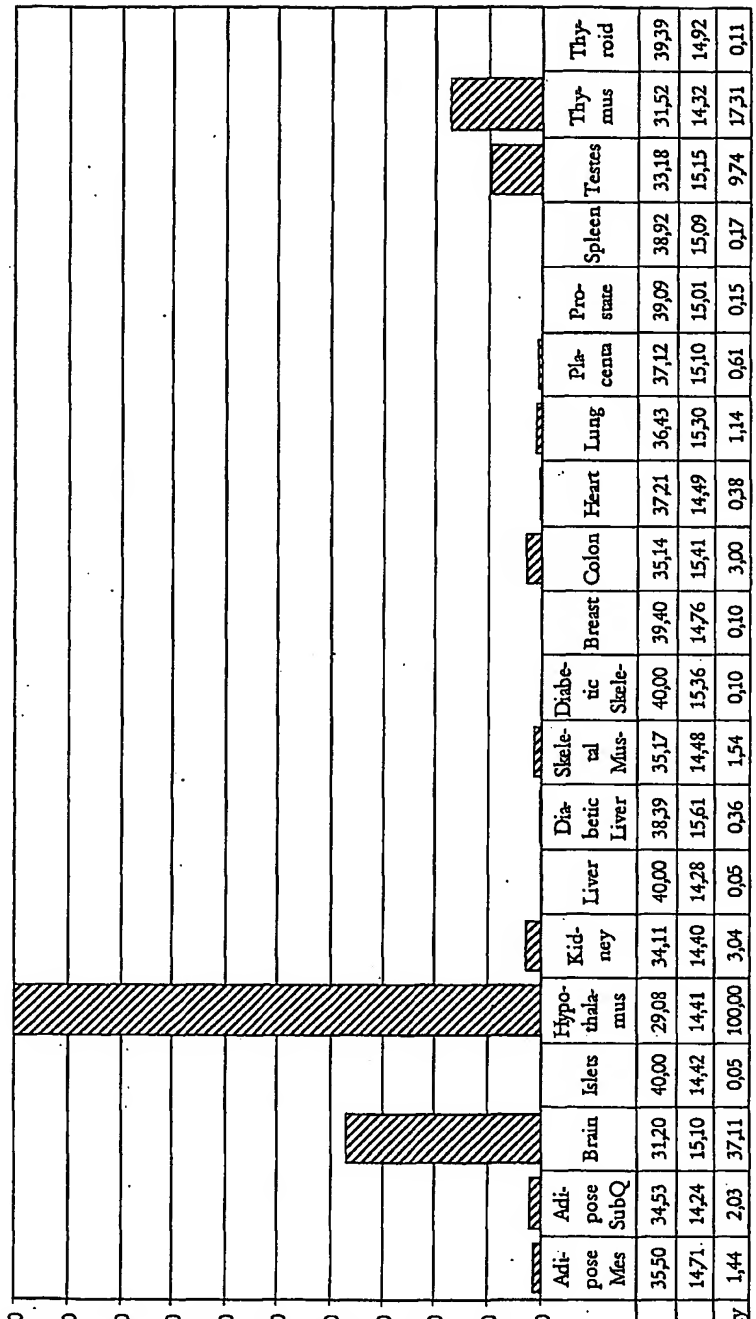
```
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        intron  1916746 1928106 0.00
+       .       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        cds     1928107 1928282 0.00
+       2       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        intron  1928283 1935529 0.00
+       .       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        cds     1935530 1935631 0.00
+       0       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        intron  1935632 1951576 0.00
+       .       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        cds     1951577 1951719 0.00
+       0       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        intron  1951720 1952940 0.00
+       .       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        cds     1952941 1953034 0.00
+       1       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        intron  1953035 1960436 0.00
+       .       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        cds     1960437 1960640 0.00
+       0       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        intron  1960641 1962833 0.00
+       .       AF086823_1
```

Fig. 16 (continued)

```
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        cds     1962834 1962987 0.00
+       0       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        intron  1962988 1982417 0.00
+       .       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        cds     1982418 1982601 0.00
+       2       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        intron  1982602 2000741 0.00
+       .       AF086823_1
gi|13653116|ref|NT_009775.3|Hs12_9932   GeneWise        cds     2000742 2000931 0.00
+       1       AF086823_1FIG. 10
```

Fig. 17

LBRI-544 Relative Expression




| | Adipose Mes | Adipose SubQ | Brain | Islets | Hypothalamus | Kidney | Liver | Diabetic Liver | Skeletal Muscle | Diabetic Skeletal | Breast | Colon | Heart | Lung | Placenta | Prostate | Spleen | Testes | Thymus | Thyroid |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 544-Ct | 35,50 | 34,53 | 31,20 | 40,00 | 29,08 | 34,11 | 40,00 | 38,39 | 35,17 | 40,00 | 39,40 | 35,14 | 37,21 | 36,43 | 37,12 | 39,09 | 38,92 | 33,18 | 31,52 | 39,39 |
| 18s-Ct | 14,71 | 14,24 | 15,10 | 14,42 | 14,41 | 14,40 | 14,28 | 15,61 | 14,48 | 15,36 | 14,76 | 15,41 | 14,49 | 15,30 | 15,10 | 15,01 | 15,09 | 15,15 | 14,32 | 14,92 |
| Relative Intensity | 1,44 | 2,03 | 37,11 | 0,05 | 100,00 | 3,04 | 0,05 | 0,36 | 1,54 | 0,10 | 0,10 | 3,00 | 0,38 | 1,14 | 0,61 | 0,15 | 0,17 | 9,74 | 17,31 | 0,11 |

Fig. 18

ATGTTGAAGTTCAAATATGGAGCGCGGAATCCTTTGGATGCTGGTGCTGC
TGAACCCATTGCCAGCCGGGCCTCCAGGCTGAATCTGTTCTTCCAGGGGA
AACCACCCTTTATGACTCAACAGCAGATGTCTCCTCTTTCCCGAGAAGGG
ATATTAGATGCCCTCTTTGTTCTCTTTGAAGAATGCAGTCAGCCTGCTCT
GATGAAGATTAAGCACGTGAGCAACTTTGTCCGGAAGTATTCCGACACCA
TAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAAGGACTTCGAAGTCAGA
AGTCTTGTAGGTTGTGGTCACTTTGCTGAAGTGCAGGTGGTAAGAGAGAA
AGCAACCGGGGACATCTATGCTATGAAAGTGATGAAGAAGAAGGCTTTAT
TGGCCCAGGAGCAGGTTTCATTTTTTGAGGAAGAGCGGAACATATTATCT
CGAAGCACAAGCCCGTGGATCCCCCAATTACAGTATGCCTTTCAGGACAA
AAATCACCTTTATCTGGTCATGGAATATCAGCCTGGAGGGGACTTGCTGT
CACTTTTGAATAGATATGAGGACCAGTTAGATGAAAACCTGATACAGTTT
TACCTAGCTGAGCTGATTTTGGCTGTTCACAGCGTTCATCTGATGGGATA
CGTGCATCGAGACATCAAGCCTGAGAACATTCTCGTTGACCGCACAGGAC
ACATCAAGCTGGTGGATTTTGGATCTGCCGCGAAAATGAATTCAAACAAG
ATGGTGAATGCCAAACTCCCGATTGGGACCCCAGATTACATGGCTCCTGA
AGTGCTGACTGTGATGAACGGGGATGGAAAAGGCACCTACGGCCTGGACT
GTGACTGGTGGTCAGTGGGCGTGATTGCCTATGAGATGATTTATGGGAGA
TCCCCCTTCGCAGAGGGAACCTCTGCCAGAACCTTCAATAACATTATGAA
TTTCCAGCGGTTTTTGAAATTTCCAGATGACCCCAAAGTGAGCAGTGACT
TTCTTGATCTGATTCAAAGCTTGTTGTGCGGCCAGAAAGAGAGACTGAAG
TTTGAAGGTCTTTGCTGCCATCCTTTCTTCTCTAAAATTGACTGGAACAA
CATTCGTAACTCTCCTCCCCCCTTCGTTCCCACCCTCAAGTCTGACGATG

Fig. 18 (continued)

ACACCTCCAATTTTGATGAACCAGAGAAGAATTCGTGGGTTTCATCCTCT
CCGTGCCAGCTGAGCCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGT
GGGGTTTTCGTACAGCAAGGCACTGGGGATTCTTGGTAGATCTGAGTCTG
TTGTGTCGGGTCTGGACTCCCCTGCCAAGACTAGCTCCATGGAAAAGAAA
CTTCTCATCAAAAGCAAAGAGCTACAAGACTCTCAGGACAAGTGTCACAA
GGTATTTATTTCCGCAGCCGGCCTCCTTCCTTGCTCCAGGATCCTCCCGT
CCGTATATGCCAAGGGATCCGCCCGGGGCCGCTGCTGGCTCTGAGCCGCC
TGATCCGTAGAGAGTGAGGCGCTCCTGCCTTCGCTGAAGTCGCGCCTCCA
GCAGCTCAGAGGGAGATGAATTCGGGCCTTGCTGTTGCTGTAAATCCTTT
AAATCTAAACCAGAGGAGGCCCTGGATTTAAACAGTCCGTTTCTCAGCAT
GACCCAGCCAGATGTCTGCTTCTTCCGGCAGGTGGCCTGGGTCCTCACCT
GTGGCTGAGATACATCCCATCTGCTTTGAGTGATGCGAAGTCTCTCTTCC
TAGTCTTTTAAAACT

Fig. 19

MLKFKYGARNPLDAGAAEPIASRASRLNLFFQGKPPFMTQQQMSPLSREG
ILDALFVLFEECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVR
SLVGCGHFAEVQVVREKATGDIYAMKVMKKKALLAQEQVSFFEEERNILS
RSTSPWIPQLQYAFQDKNHLYLVMEYQPGGDLLSLLNRYEDQLDENLIQF
YLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVDFGSAAKMNSNK
MVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYGR
SPFAEGTSARTFNNIMNFQRFLKFPDDPKVSSDFLDLIQSLLCGQKERLK
FEGLCCHPFFSKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSSS
PCQLSPSGFSGEELPFVGFSYSKALGILGRSESVVSGLDSPAKTSSMEKK
LLIKSKELQDSQDKCHKVFISAAGLLPCSRILPSVYAKGSARGRCWL

Fig. 20

TBLASTN - alignment of 544_Protein against emnew|AX166510|AX166510
Sequence 1 from Patent WO0138503.//:gbnew|AX166510|AX166510 Sequence 1 from Patent WO0138503.

This hit is scoring at : 0.0 (expectation value)
Alignment length (overlap) : 469
Identities : 99 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Hit reading frame : +1
Database searched : nrnee_1_;

```
Q:      1 MLKFKYGARNPLDAGAAEPIASRASRLNLFFQGKPPFMTQQQMSPLSREGILDALFVLFE
          MLKFKYGARNPLDAGAAEPIASRASRLNLFFQGKPPFMTQQQMSPLSREGILDALFVLFE
H:      1 MLKFKYGARNPLDAGAAEPIASRASRLNLFFQGKPPFMTQQQMSPLSREGILDALFVLFE

          ECSQPALMKIKHVSNFVRK-YSDTIAELQELQPSAKDFEVRSLVGCGHFAEVQVVREKAT
          ECSQPALMKIKHVSNFV : YSDTIAELQELQPSAKDFEVRSLVGCGHFAEVQVVREKAT
          ECSQPALMKIKHVSNFVPEVYSDTIAELQELQPSAKDFEVRSLVGCGHFAEVQVVREKAT

          GDIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLYLVMEYQPG
          GDIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLYLVMEYQPG
          GDIYAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLYLVMEYQPG
```

Fig. 20 (continued)

```
GDLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVD
GDLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVD
GDLLSLLNRYEDQLDENLIQFYLAELILAVHSVHLMGYVHRDIKPENILVDRTGHIKLVD

FGSAAKMNSNKMVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYG
FGSAAKMNSNKMVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYG
FGSAAKMNSNKMVNAKLPIGTPDYMAPEVLTVMNGDGKGTYGLDCDWWSVGVIAYEMIYG

RSPFAEGTSARTFNNIMNFQRFLKFPDDPKVSSDFLDLIQSLLCGQKERLKFEGLCCHPF
RSPFAEGTSARTFNNIMNFQRFLKFPDDPKVSSDFLDLIQSLLCGQKERLKFEGLCCHPF
RSPFAEGTSARTFNNIMNFQRFLKFPDDPKVSSDFLDLIQSLLCGQKERLKFEGLCCHPF

FSKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSSSPCQLSPSGFSGEELPFVGF
FSKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSSSPCQLSPSGFSGEELPFVGF
FSKIDWNNIRNSPPPFVPTLKSDDDTSNFDEPEKNSWVSSSPCQLSPSGFSGEELPFVGF

SYSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKV      468
SYSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHK:
SYSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKM     1407
```

SEQUENCE LISTING

<110> Bayer AG

<120> REGULATION OF HUMAN CITRON RHO/RAC-INTERACTING KINASE-SHORT KINASE

<130> LIO371 Foreign Countries

<150> US 60/301,853

<151> 2001-07-02

<150> US 60/337,130

<151> 2001-12-10

<150> US 60/375,015

<151> 2002-04-25

<160> 9

<170> PatentIn version 3.1

<210> 1

<211> 1485

<212> DNA

<213> Homo sapiens

<400> 1

atgttgaagt tcaaatatgg agcgcggaat cctttggatg ctggtgctgc tgaacccatt 60
gccagccggg cctccaggct gaatctgttc ttccagggga aaccaccctt tatgactcaa 120
cagcagatgt ctcctctttc ccgagaaggg atattagatg ccctctttgt tctctttgaa 180
gaatgcagtc agcctgctct gatgaagatt aagcacgtga gcaactttgt ccggaagtat 240
tccgacacca tagctgagtt acaggagctc cagccttcgg caaaggactt cgaagtcaga 300
agtcttgtag gttgtggtca ctttgctgaa gtgcaggtgg taagagagaa agcaaccggg 360
gacatctatg ctatgaaagt gatgaagaag aaggctttat tggcccagga gcaggtttca 420
ttttttgagg aagagcggaa catattatct cgaagcacaa gcccgtggat cccccaatta 480
cagtatgcct ttcaggacaa aaatcacctt tatctggtca tggaatatca gcctggaggg 540

```
gacttgctgt cacttttgaa tagatatgag gaccagttag atgaaaacct gatacagttt   600
tacctagctg agctgatttt ggctgttcac agcgttcatc tgatgggata cgtgcatcga   660
gacatcaagc ctgagaacat tctcgttgac cgcacaggac acatcaagct ggtggatttt   720
ggatctgccg cgaaaatgaa ttcaaacaag atggtgaatg ccaaactccc gattgggacc   780
ccagattaca tggctcctga agtgctgact gtgatgaacg gggatggaaa aggcacctac   840
ggcctggact gtgactggtg gtcagtgggc gtgattgcct atgagatgat ttatgggaga   900
tccccttcg cagagggaac ctctgccaga accttcaata acattatgaa tttccagcgg   960
tttttgaaat ttccagatga ccccaaagtg agcagtgact ttcttgatct gattcaaagc  1020
ttgttgtgcg gccagaaaga gagactgaag tttgaaggtc tttgctgcca tcctttcttc  1080
tctaaaattg actggaacaa cattcgtaac tctcctcccc ccttcgttcc caccctcaag  1140
tctgacgatg acacctccaa ttttgatgaa ccagagaaga attcgtgggt ttcatcctct  1200
ccgtgccagc tgagcccctc aggcttctcg ggtgaagaac tgccgtttgt ggggttttcg  1260
tacagcaagg cactggggat tcttggtaga tctgagtctg ttgtgtcggg tctggactcc  1320
cctgccaaga ctagctccat ggaaaagaaa cttctcatca aaagcaaaga gctacaagac  1380
tctcaggaca agtgtcacaa ggtatttatt tccgcagccg gcctccttcc ttgctccagg  1440
atcctcccgt ccgtatatgc caagggatcc gcccggggcc gctgc                  1485
```

<210> 2

<211> 495

<212> PRT

<213> Homo sapiens

<400> 2

```
Met Leu Lys Phe Lys Tyr Gly Ala Arg Asn Pro Leu Asp Ala Gly Ala
1               5                  10                  15
Ala Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
            20                  25                  30
Gly Lys Pro Pro Phe Met Thr Gln Gln Gln Met Ser Pro Leu Ser Arg
        35                  40                  45
Glu Gly Ile Leu Asp Ala Leu Phe Val Leu Phe Glu Glu Cys Ser Gln
    50                  55                  60
Pro Ala Leu Met Lys Ile Lys His Val Ser Asn Phe Val Arg Lys Tyr
65                  70                  75                  80
Ser Asp Thr Ile Ala Glu Leu Gln Glu Leu Gln Pro Ser Ala Lys Asp
                85                  90                  95
Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
            100                 105                 110
Val Val Arg Glu Lys Ala Thr Gly Asp Ile Tyr Ala Met Lys Val Met
        115                 120                 125
Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
    130                 135                 140
Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
145                 150                 155                 160
Gln Tyr Ala Phe Gln Asp Lys Asn His Leu Tyr Leu Val Met Glu Tyr
                165                 170                 175
Gln Pro Gly Gly Asp Leu Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
            180                 185                 190
Leu Asp Glu Asn Leu Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
        195                 200                 205
Val His Ser Val His Leu Met Gly Tyr Val His Arg Asp Ile Lys Pro
    210                 215                 220
Glu Asn Ile Leu Val Asp Arg Thr Gly His Ile Lys Leu Val Asp Phe
225                 230                 235                 240
Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Met Val Asn Ala Lys Leu
                245                 250                 255
Pro Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met
            260                 265                 270
```

```
Asn Gly Asp Gly Lys Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser
        275                 280                 285
Val Gly Val Ile Ala Tyr Glu Met Ile Tyr Gly Arg Ser Pro Phe Ala
    290                 295                 300
Glu Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg
305                 310                 315                 320
Phe Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Asp Phe Leu Asp
                325                 330                 335
Leu Ile Gln Ser Leu Leu Cys Gly Gln Lys Glu Arg Leu Lys Phe Glu
            340                 345                 350
Gly Leu Cys Cys His Pro Phe Phe Ser Lys Ile Asp Trp Asn Asn Ile
        355                 360                 365
Arg Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp
    370                 375                 380
Thr Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Val Ser Ser Ser
385                 390                 395                 400
Pro Cys Gln Leu Ser Pro Ser Gly Phe Ser Gly Glu Glu Leu Pro Phe
                405                 410                 415
Val Gly Phe Ser Tyr Ser Lys Ala Leu Gly Ile Leu Gly Arg Ser Glu
            420                 425                 430
Ser Val Val Ser Gly Leu Asp Ser Pro Ala Lys Thr Ser Ser Met Glu
        435                 440                 445
Lys Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys
    450                 455                 460
Cys His Lys Val Phe Ile Ser Ala Ala Gly Leu Leu Pro Cys Ser Arg
465                 470                 475                 480
Ile Leu Pro Ser Val Tyr Ala Lys Gly Ser Ala Arg Gly Arg Cys
                485                 490                 495
```

<210> 3

<211> 494

<212> PRT

<213> Homo sapiens

<400> 3

```
Met Leu Lys Phe Lys Tyr Gly Val Arg Asn Pro Pro Glu Ala Ser Ala
1               5                   10                  15
Ser Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
            20                  25                  30
Gly Lys Pro Pro Leu Met Thr Gln Gln Gln Met Ser Ala Leu Ser Arg
        35                  40                  45
Glu Gly Met Leu Asp Ala Leu Phe Ala Leu Phe Glu Glu Cys Ser Gln
    50                  55                  60
Pro Ala Leu Met Lys Met Lys His Val Ser Ser Phe Val Gln Lys Tyr
65                  70                  75                  80
Ser Asp Thr Ile Ala Glu Leu Arg Glu Leu Gln Pro Ser Ala Arg Asp
                85                  90                  95
Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
            100                 105                 110
Val Val Arg Glu Lys Ala Thr Gly Asp Val Tyr Ala Met Lys Ile Met
        115                 120                 125
Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
    130                 135                 140
Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
145                 150                 155                 160
Gln Tyr Ala Phe Gln Asp Lys Asn Asn Leu Tyr Leu Val Met Glu Tyr
                165                 170                 175
```

```
Gln Pro Gly Gly Asp Phe Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
            180                 185                 190
Leu Asp Glu Ser Met Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
        195                 200                 205
Val His Ser Val His Gln Met Gly Tyr Val His Arg Asp Ile Lys Pro
    210                 215                 220
Glu Asn Ile Leu Ile Asp Arg Thr Gly Glu Ile Lys Leu Val Asp Phe
225                 230                 235                 240
Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Val Asp Ala Lys Leu Pro
                245                 250                 255
Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met Asn
            260                 265                 270
Glu Asp Arg Arg Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser Val
        275                 280                 285
Gly Val Val Ala Tyr Glu Met Val Tyr Gly Lys Thr Pro Phe Thr Glu
    290                 295                 300
Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg Phe
305                 310                 315                 320
Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Glu Leu Leu Asp Leu
                325                 330                 335
Leu Gln Ser Leu Leu Cys Val Gln Lys Glu Arg Leu Lys Phe Glu Gly
            340                 345                 350
Leu Cys Cys His Pro Phe Phe Ala Arg Thr Asp Trp Asn Asn Ile Arg
        355                 360                 365
Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp Thr
    370                 375                 380
Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Ala Phe Ile Leu Cys
385                 390                 395                 400
Val Pro Ala Glu Pro Leu Ala Phe Ser Gly Glu Glu Leu Pro Phe Val
                405                 410                 415
Gly Phe Ser Tyr Ser Lys Ala Leu Gly Tyr Leu Gly Arg Ser Glu Ser
            420                 425                 430
Val Val Ser Ser Leu Asp Ser Pro Ala Lys Val Ser Ser Met Glu Lys
        435                 440                 445
Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys Cys
    450                 455                 460
His Lys Val Ser Ile Ser Thr Ala Gly Leu Arg Pro Cys Ser Arg Ile
465                 470                 475                 480
Leu Gln Ser Ile Tyr Ala Glu Gly Ser Ala Gly Gly His Cys
                485                 490
<210>  4

<211>  2055

<212>  PRT

<213>  Homo sapiens


<400>  4
Met Leu Lys Phe Lys Tyr Gly Val Arg Asn Pro Pro Glu Ala Ser Ala
1                 5                  10                  15
Ser Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
            20                  25                  30
Gly Lys Pro Pro Leu Met Thr Gln Gln Gln Met Ser Ala Leu Ser Arg
        35                  40                  45
Glu Gly Met Leu Asp Ala Leu Phe Ala Leu Phe Glu Glu Cys Ser Gln
    50                  55                  60
Pro Ala Leu Met Lys Met Lys His Val Ser Ser Phe Val Gln Lys Tyr
65                  70                  75                  80
```

```
Ser Asp Thr Ile Ala Glu Leu Arg Glu Leu Gln Pro Ser Ala Arg Asp
                85                  90                  95
Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
            100                 105                 110
Val Val Arg Glu Lys Ala Thr Gly Asp Val Tyr Ala Met Lys Ile Met
        115                 120                 125
Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
    130                 135                 140
Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
145                 150                 155                 160
Gln Tyr Ala Phe Gln Asp Lys Asn Asn Leu Tyr Leu Val Met Glu Tyr
                165                 170                 175
Gln Pro Gly Gly Asp Phe Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
            180                 185                 190
Leu Asp Glu Ser Met Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
        195                 200                 205
Val His Ser Val His Gln Met Gly Tyr Val His Arg Asp Ile Lys Pro
    210                 215                 220
Glu Asn Ile Leu Ile Asp Arg Thr Gly Glu Ile Lys Leu Val Asp Phe
225                 230                 235                 240
Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Val Asp Ala Lys Leu Pro
                245                 250                 255
Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met Asn
            260                 265                 270
Glu Asp Arg Arg Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser Val
        275                 280                 285
Gly Val Val Ala Tyr Glu Met Val Tyr Gly Lys Thr Pro Phe Thr Glu
    290                 295                 300
Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg Phe
305                 310                 315                 320
Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Glu Leu Leu Asp Leu
                325                 330                 335
Leu Gln Ser Leu Leu Cys Val Gln Lys Glu Arg Leu Lys Phe Glu Gly
            340                 345                 350
Leu Cys Cys His Pro Phe Phe Ala Arg Thr Asp Trp Asn Asn Ile Arg
        355                 360                 365
Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp Thr
    370                 375                 380
Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Ala Phe Ile Leu Cys
385                 390                 395                 400
Val Pro Ala Glu Pro Leu Ala Phe Ser Gly Glu Glu Leu Pro Phe Val
                405                 410                 415
Gly Phe Ser Tyr Ser Lys Ala Leu Gly Tyr Leu Gly Arg Ser Glu Ser
            420                 425                 430
Val Val Ser Ser Leu Asp Ser Pro Ala Lys Val Ser Ser Met Glu Lys
        435                 440                 445
Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys Cys
    450                 455                 460
His Lys Met Glu Gln Glu Met Thr Arg Leu His Arg Arg Val Ser Glu
465                 470                 475                 480
Val Glu Ala Val Leu Ser Gln Lys Glu Val Glu Leu Lys Ala Ser Glu
                485                 490                 495
Thr Gln Arg Ser Leu Leu Glu Gln Asp Leu Ala Thr Tyr Ile Thr Glu
            500                 505                 510
Cys Ser Ser Leu Lys Arg Ser Leu Glu Gln Ala Arg Met Glu Val Ser
        515                 520                 525
Gln Glu Asp Asp Lys Ala Leu Gln Leu Leu His Asp Ile Arg Glu Gln
    530                 535                 540
Ser Arg Lys Leu Gln Glu Ile Lys Glu Gln Glu Tyr Gln Ala Gln Val
545                 550                 555                 560
```

```
Glu Glu Met Arg Leu Met Met Asn Gln Leu Glu Glu Asp Leu Val Ser
                565                 570                 575
Ala Arg Arg Arg Ser Asp Leu Tyr Glu Ser Glu Leu Arg Glu Ser Arg
            580                 585                 590
Leu Ala Ala Glu Glu Phe Lys Arg Lys Ala Asn Glu Cys Gln His Lys
        595                 600                 605
Leu Met Lys Ala Lys Asp Gln Gly Lys Pro Glu Val Gly Glu Tyr Ser
    610                 615                 620
Lys Leu Glu Lys Ile Asn Ala Glu Gln Gln Leu Lys Ile Gln Glu Leu
625                 630                 635                 640
Gln Glu Lys Leu Glu Lys Ala Val Lys Ala Ser Thr Glu Ala Thr Glu
                645                 650                 655
Leu Leu Gln Asn Ile Arg Gln Ala Lys Glu Arg Ala Glu Arg Glu Leu
            660                 665                 670
Glu Lys Leu His Asn Arg Glu Asp Ser Ser Glu Gly Ile Lys Lys Lys
        675                 680                 685
Leu Val Glu Ala Glu Glu Arg Arg His Ser Leu Glu Asn Lys Val Lys
    690                 695                 700
Arg Leu Glu Thr Met Glu Arg Arg Glu Asn Arg Leu Lys Asp Asp Ile
705                 710                 715                 720
Gln Thr Lys Ser Glu Gln Ile Gln Gln Met Ala Asp Lys Ile Leu Glu
                725                 730                 735
Leu Glu Glu Lys His Arg Glu Ala Gln Val Ser Ala Gln His Leu Glu
            740                 745                 750
Val His Leu Lys Gln Lys Glu Gln His Tyr Glu Glu Lys Ile Lys Val
        755                 760                 765
Leu Asp Asn Gln Ile Lys Lys Asp Leu Ala Asp Lys Glu Ser Leu Glu
    770                 775                 780
Asn Met Met Gln Arg His Glu Glu Glu Ala His Glu Lys Gly Lys Ile
785                 790                 795                 800
Leu Ser Glu Gln Lys Ala Met Ile Asn Ala Met Asp Ser Lys Ile Arg
                805                 810                 815
Ser Leu Glu Gln Arg Ile Val Glu Leu Ser Glu Ala Asn Lys Leu Ala
            820                 825                 830
Ala Asn Ser Ser Leu Phe Thr Gln Arg Asn Met Lys Ala Gln Glu Glu
        835                 840                 845
Met Ile Ser Glu Leu Arg Gln Gln Lys Phe Tyr Leu Glu Thr Gln Ala
    850                 855                 860
Gly Lys Leu Glu Ala Gln Asn Arg Lys Leu Glu Glu Gln Leu Glu Lys
865                 870                 875                 880
Ile Ser His Gln Asp His Ser Asp Lys Ser Arg Leu Leu Glu Leu Glu
                885                 890                 895
Thr Arg Leu Arg Glu Val Ser Leu Glu His Glu Glu Gln Lys Leu Glu
            900                 905                 910
Leu Lys Arg Gln Leu Thr Glu Leu Gln Leu Ser Leu Gln Glu Arg Glu
        915                 920                 925
Ser Gln Leu Thr Ala Leu Gln Ala Ala Arg Ala Ala Leu Glu Ser Gln
    930                 935                 940
Leu Arg Gln Ala Lys Thr Glu Leu Glu Glu Thr Thr Ala Glu Ala Glu
945                 950                 955                 960
Glu Glu Ile Gln Ala Leu Thr Ala His Arg Asp Glu Ile Gln Arg Lys
                965                 970                 975
Phe Asp Ala Leu Arg Asn Ser Cys Thr Val Ile Thr Asp Leu Glu Glu
            980                 985                 990
Gln Leu Asn Gln Leu Thr Glu Asp Asn Ala Glu Leu Asn Asn Gln Asn
        995                 1000                1005
Phe Tyr Leu Ser Lys Gln Leu Asp Glu Ala Ser Gly Ala Asn Asp
    1010                1015                1020
Glu Ile Val Gln Leu Arg Ser Glu Val Asp His Leu Arg Arg Glu
    1025                1030                1035
Ile Thr Glu Arg Glu Met Gln Leu Thr Ser Gln Lys Gln Thr Met
    1040                1045                1050
```

```
Glu Ala Leu Lys Thr Thr Cys Thr Met Leu Glu Glu Gln Val Leu
1055                1060                1065
Asp Leu Glu Ala Leu Asn Asp Glu Leu Leu Glu Lys Glu Arg Gln
1070                1075                1080
Trp Glu Ala Trp Arg Ser Val Leu Gly Asp Glu Lys Ser Gln Phe
1085                1090                1095
Glu Cys Arg Val Arg Glu Leu Gln Arg Met Leu Asp Thr Glu Lys
1100                1105                1110
Gln Ser Arg Ala Arg Ala Asp Gln Arg Ile Thr Glu Ser Arg Gln
1115                1120                1125
Val Val Glu Leu Ala Val Lys Glu His Lys Ala Glu Ile Leu Ala
1130                1135                1140
Leu Gln Gln Ala Leu Lys Glu Gln Lys Leu Lys Ala Glu Ser Leu
1145                1150                1155
Ser Asp Lys Leu Asn Asp Leu Glu Lys Lys His Ala Met Leu Glu
1160                1165                1170
Met Asn Ala Arg Ser Leu Gln Gln Lys Leu Glu Thr Glu Arg Glu
1175                1180                1185
Leu Lys Gln Arg Leu Leu Glu Glu Gln Ala Lys Leu Gln Gln Gln
1190                1195                1200
Met Asp Leu Gln Lys Asn His Ile Phe Arg Leu Thr Gln Gly Leu
1205                1210                1215
Gln Glu Ala Leu Asp Arg Ala Asp Leu Leu Lys Thr Glu Arg Ser
1220                1225                1230
Asp Leu Glu Tyr Gln Leu Glu Asn Ile Gln Val Leu Tyr Ser His
1235                1240                1245
Glu Lys Val Lys Met Glu Gly Thr Ile Ser Gln Gln Thr Lys Leu
1250                1255                1260
Ile Asp Phe Leu Gln Ala Lys Met Asp Gln Pro Ala Lys Lys Lys
1265                1270                1275
Lys Val Pro Leu Gln Tyr Asn Glu Leu Lys Leu Ala Leu Glu Lys
1280                1285                1290
Glu Lys Ala Arg Cys Ala Glu Leu Glu Glu Ala Leu Gln Lys Thr
1295                1300                1305
Arg Ile Glu Leu Arg Ser Ala Arg Glu Glu Ala Ala His Arg Lys
1310                1315                1320
Ala Thr Asp His Pro His Pro Ser Thr Pro Ala Thr Ala Arg Gln
1325                1330                1335
Gln Ile Ala Met Ser Ala Ile Val Arg Ser Pro Glu His Gln Pro
1340                1345                1350
Ser Ala Met Ser Leu Leu Ala Pro Pro Ser Ser Arg Arg Lys Glu
1355                1360                1365
Ser Ser Thr Pro Glu Glu Phe Ser Arg Arg Leu Lys Glu Arg Met
1370                1375                1380
His His Asn Ile Pro His Arg Phe Asn Val Gly Leu Asn Met Arg
1385                1390                1395
Ala Thr Lys Cys Ala Val Cys Leu Asp Thr Val His Phe Gly Arg
1400                1405                1410
Gln Ala Ser Lys Cys Leu Glu Cys Gln Val Met Cys His Pro Lys
1415                1420                1425
Cys Ser Thr Cys Leu Pro Ala Thr Cys Gly Leu Pro Ala Glu Tyr
1430                1435                1440
Ala Thr His Phe Thr Glu Ala Phe Cys Arg Asp Lys Met Asn Ser
1445                1450                1455
Pro Gly Leu Gln Ser Lys Glu Pro Gly Ser Ser Leu His Leu Glu
1460                1465                1470
Gly Trp Met Lys Val Pro Arg Asn Asn Lys Arg Gly Gln Gln Gly
1475                1480                1485
Trp Asp Arg Lys Tyr Ile Val Leu Glu Gly Ser Lys Val Leu Ile
1490                1495                1500
Tyr Asp Asn Glu Ala Arg Glu Ala Gly Gln Arg Pro Val Glu Glu
1505                1510                1515
```

Phe Glu Leu Cys Leu Pro Asp Gly Asp Val Ser Ile His Gly Ala
1520 1525 1530
Val Gly Ala Ser Glu Leu Ala Asn Thr Ala Lys Ala Asp Val Pro
1535 1540 1545
Tyr Ile Leu Lys Met Glu Ser His Pro His Thr Thr Cys Trp Pro
1550 1555 1560
Gly Arg Thr Leu Tyr Leu Leu Ala Pro Ser Phe Pro Asp Lys Gln
1565 1570 1575
Arg Trp Val Thr Ala Leu Glu Ser Val Val Ala Gly Gly Arg Val
1580 1585 1590
Ser Arg Glu Lys Ala Glu Ala Asp Ala Lys Leu Leu Gly Asn Ser
1595 1600 1605
Leu Leu Lys Leu Glu Gly Asp Asp Arg Leu Asp Met Asn Cys Thr
1610 1615 1620
Leu Pro Phe Ser Asp Gln Val Val Leu Val Gly Thr Glu Glu Gly
1625 1630 1635
Leu Tyr Ala Leu Asn Val Leu Lys Asn Ser Leu Thr His Ile Pro
1640 1645 1650
Gly Ile Gly Ala Val Phe Gln Ile Tyr Ile Ile Lys Asp Leu Glu
1655 1660 1665
Lys Leu Leu Met Ile Ala Gly Glu Glu Arg Ala Leu Cys Leu Val
1670 1675 1680
Asp Val Lys Lys Val Lys Gln Ser Leu Ala Gln Ser His Leu Pro
1685 1690 1695
Ala Gln Pro Asp Val Ser Pro Asn Ile Phe Glu Ala Val Lys Gly
1700 1705 1710
Cys His Leu Phe Ala Ala Gly Lys Ile Glu Asn Ser Leu Cys Ile
1715 1720 1725
Cys Ala Ala Met Pro Ser Lys Val Val Ile Leu Arg Tyr Asn Asp
1730 1735 1740
Asn Leu Ser Lys Tyr Cys Ile Arg Lys Glu Ile Glu Thr Ser Glu
1745 1750 1755
Pro Cys Ser Cys Ile His Phe Thr Asn Tyr Ser Ile Leu Ile Gly
1760 1765 1770
Thr Asn Lys Phe Tyr Glu Ile Asp Met Lys Gln Tyr Thr Leu Asp
1775 1780 1785
Glu Phe Leu Asp Lys Asn Asp His Ser Leu Ala Pro Ala Val Phe
1790 1795 1800
Ala Ser Ser Ser Asn Ser Phe Pro Val Ser Ile Val Gln Ala Asn
1805 1810 1815
Ser Ala Gly Gln Arg Glu Glu Tyr Leu Leu Cys Phe His Glu Phe
1820 1825 1830
Gly Val Phe Val Asp Ser Tyr Gly Arg Arg Ser Arg Thr Asp Asp
1835 1840 1845
Leu Lys Trp Ser Arg Leu Pro Leu Ala Phe Ala Tyr Arg Glu Pro
1850 1855 1860
Tyr Leu Phe Val Thr His Phe Asn Ser Leu Glu Val Ile Glu Ile
1865 1870 1875
Gln Ala Arg Ser Ser Leu Gly Ser Pro Ala Arg Ala Tyr Leu Glu
1880 1885 1890
Ile Pro Asn Pro Arg Tyr Leu Gly Pro Ala Ile Ser Ser Gly Ala
1895 1900 1905
Ile Tyr Leu Ala Ser Ser Tyr Gln Asp Lys Leu Arg Val Ile Cys
1910 1915 1920
Cys Lys Gly Asn Leu Val Lys Glu Ser Gly Thr Glu Gln His Arg
1925 1930 1935
Val Pro Ser Thr Ser Arg Ser Ser Pro Asn Lys Arg Gly Pro Pro
1940 1945 1950
Thr Tyr Asn Glu His Ile Thr Lys Arg Val Ala Ser Ser Pro Ala
1955 1960 1965
Pro Pro Glu Gly Pro Ser His Pro Arg Glu Pro Ser Thr Pro His
1970 1975 1980

Arg Tyr Arg Asp Arg Glu Gly Arg Thr Glu Leu Arg Arg Asp Lys
1985 1990 1995
Ser Pro Gly Arg Pro Leu Glu Arg Glu Lys Ser Pro Gly Arg Met
2000 2005 2010
Leu Ser Thr Arg Arg Glu Arg Ser Pro Gly Arg Leu Phe Glu Asp
2015 2020 2025
Ser Ser Arg Gly Arg Leu Pro Ala Gly Ala Val Arg Thr Pro Leu
2030 2035 2040
Ser Gln Val Asn Lys Val Trp Asp Gln Ser Ser Val
2045 2050 2055

<210> 5

<211> 6780

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (6244)..(6245)

<223> n = a, t, g, or c

<220>

<221> misc_feature

<222> (6259)..(6259)

<223> n = a, t, g, or c

<220>

<221> misc_feature

<222> (6269)..(6269)

<223> n = a, t, g, or c

<220>

<221> misc_feature

<222> (6266)..(6266)

<223> n = a, t, g, or c

<400> 5

gggcggggct gagggcggcg ggggcgggcc gcccgagctg ggagggcggc ggcgccgagg 60
ggaggagagc ggcccatgga cccgcggggc ccggcgcccc agactctgcg ccgtcgggac 120
ggagcccaag atgtcggcct aggccggggc gcgacgacgc ggacggggcg gcgaggaggc 180
gccgctgctg ccggggctcg cagccgccga gcccccgagg gcgcgccctg acggactggc 240
cgagccggcg gtgagaggcc ggcgcgtctg gagcgggccg cgcggcacca tgtcggccaa 300

ggtgcggctc aagaagctgg agcagctgct cctggacggg ccctggcgca acgagagcgc 360
cctgagcgtg gaaacgctgc tcgacgtgct cgtctgcctg tacaccgagt gcagccactc 420
ggccctgcgc cgcgacaagt acgtggccga gttcctcgag tgggctaaac catttacaca 480
gctggtgaaa gaaatgcagc ttcatcgaga agactttgaa ataattaaag taattggaag 540
aggtgctttt ggtgaggttg ctgttgtcaa aatgaagaat actgaacgaa tttatgcaat 600
gaaaatcctc aacaagtggg agatgctgaa aagagcagag accgcgtgct tccgagagga 660
gcgcgatgtg ctggtgaacg gcgactgcca gtggatcacc gcgctgcact acgcctttca 720
ggacgagaac cacctgtact tagtcatgga ttactatgtg ggtggtgatt tactgaccct 780
gctcagcaaa tttgaagaca agcttccgga agatatggcg aggttctaca ttggtgaaat 840
ggtgctggcc attgactcca tccatcagct tcattacgtg cacagagaca ttaaacctga 900
caatgtcctt ttggacgtga atggtcatat ccgcctggct gactttggat catgtttgaa 960
gatgaatgat gatggcactg tgcagtcctc cgtggccgtg ggcacacctg actacatctc 1020
gccggagatc ctgcaggcga tggaggacgg catgggcaaa tacgggcctg agtgtgactg 1080
gtggtctctg ggtgtctgca tgtatgagat gctctatgga gaaacgccgt tttatgcgga 1140
gtcactcgtg gagacctatg ggaagatcat gaaccatgaa gagcgattcc agttcccatc 1200
ccatgtcacg gatgtatctg aagaagcgaa ggacctcatc cagagactga tctgcagtag 1260
agaacgccgg ctggggcaga atggaataga ggatttcaaa aagcatgcgt tttttgaagg 1320
tctaaattgg gaaaatatac gaaacctaga agcaccttat attcctgatg tgagcagtcc 1380
ctctgacaca tccaacttcg acgtggatga cgacgtgctg agaaacacgg aaatattacc 1440
tcctggttct cacacaggct tttctggatt acatttgcca ttcattggtt ttacattcac 1500
aacggaaagc tgtttttctg atcgaggctc tctgaagagc ataatgcagt ccaacacatt 1560
aaccaaagat gaggatgtgc agcgggacct ggagcacagc ctgcagatgg aagcttacga 1620
gaggaggatt cggaggctgg aacaggagaa gctggagctg agcaggaagc tgcaagagtc 1680
cacccagacc gtgcagtccc tccacggctc atctcgggcc ctcagcaatt caaaccgaga 1740
taaagaaatc aaaaagctaa atgaagaaat cgaacgcttg aagaataaaa tagcagattc 1800
aaacaggctg gagcgacagc ttgaggacac agtggcgctt cgccaagagc gtgaggactc 1860
cacgcagcgg ctgcgggggc tggagaagca gcaccgcgtg gtccggcagg agaaggagga 1920
gctgcacaag caactggttg aagcctcaga gcggttgaaa tcccaggcca aggaactcaa 1980
agatgcccat cagcagcgaa agctggccct gcaggagttc tcggagctga acgagcgcat 2040
ggcagagctc cgtgcccaga agcagaaggt gtcccggcag ctgcgagaca aggaggagga 2100
gatggaggtg gccacgcaga aggtggacgc catgcggcag gaaatgcgga gagctgagaa 2160
gctcaggaaa gagctggaag ctcagcttga tgatgctgtt gctgaggcct ccaaggagcg 2220
caagcttcgt gagcacagcg agaacttctg caagcaaatg gaaagcgagc tggaggccct 2280
caaggtgaag caaggaggcc ggggagcggg tgccacctta gagcaccagc aagagatttc 2340
caaaatcaaa tccgagctgg agaagaaagt cttattttat gaagaggaat tggtcagacg 2400
tgaggcctcc catgtgctag aagtgaaaaa tgtgaagaag gaggtgcatg attcagaaag 2460
ccaccagctg gccctgcaga aagaaatctt gatgttaaaa gataagttag aaaagtcaaa 2520
gcgagaacgg cataacgaga tggaggaggc agtaggtaca ataaaagata aatacgaacg 2580
agaaagagcg atgctgtttg atgaaaacaa gaagctaact gctgaaaatg aaaagctctg 2640
ttcctttgtg gataaactca cagctcaaaa tagacagctg gaggatgagc tgcaggatct 2700
ggcagccaag aaggagtcag tggcccactg ggaagctcag attgcggaaa tcattcagtg 2760
ggtcagtgac gagaaagatg cccggggtta ccttcaagct cttgcttcca agatgaccga 2820
agagctcgag gctttgagga gttctagtct ggggtcaaga acactggacc cgctgtggaa 2880
ggtgcgccgc agccagaagc tggacatgtc cgcgcggctg gagctgcagt cggccctgga 2940
ggcggagatc cgggccaagc agcttgtcca ggaggagctc aggaaggtca aggacgccaa 3000
cctcaccttg gaaagcaaac taaaggattc cgaagccaaa aacagagaat tattagaaga 3060
aatggaaatt ttgaagaaaa agatggaaga aaaattcaga gcagatactg ggctcaaact 3120
tccagatttt caggattcca tttttgagta tttcaacact gctcctcttg cacatgacct 3180
gacatttaga accagctcag ctagtgagca agaaacacaa gctccgaagc cagaagcgtc 3240
cccgtcgatg tctgtggctg catcagagca gcaggaggac atggctcggc ccccgcagag 3300
gccatccgct gtgccgttgc ccaccacgca ggccctggtt ctggctggac cgaagccaaa 3360
agctcaccag ttcagcatca agtccttctc cagccctact cagtgcagcc actgcacctc 3420
cctgatggtt gggctgatcc ggcagggcta cgcctgcgag gtgtgttcct ttgcttgcca 3480
cgtgtcctgc aaagacggtg ccccccaggt gtgcccaata cctcccgagc agtccaagag 3540
gcctctgggc gtggacgtgc agcgaggcat cggaacagcc tacaaaggcc atgtcaaggt 3600
cccaaagccc acgggggtga agaagggatg gcagcgcgca tatgcagtcg tctgtgagtg 3660
caagctcttc ctgtatgatc tgcctgaagg aaaatccacc cagcctggtg tcattgcgag 3720
ccaagtcttg gatctcagag atgacgagtt ttccgtgagc tcagtcctgg cctcagatgt 3780
cattcatgct acacgccgag atattccatg tatattcagg gtgacggcct ctctcttagg 3840
tgcaccttct aagaccagct cgctgctcat tctgacagaa aatgagaatg aaaagaggaa 3900
gtgggttggg attctagaag gactccagtc catccttcat aaaaaccggc tgaggaatca 3960
ggtcgtgcat gttcccttgg aagcctacga cagctcgctg cctctcatca aggccatcct 4020

gacagctgcc atcgtggatg cagacaggat tgcagtcggc ctagaagaag ggctctatgt 4080
catagaggtc acccgagatg tgatcgtccg tgccgctgac tgtaagaagg tacaccagat 4140
cgagcttgct cccagggaga agatcgtaat cctcctctgt ggccggaacc accatgtgca 4200
cctctatccg tggtcgtccc ttgatggagc ggaaggcagc tttgacatca agcttccgga 4260
aaccaaaggc tgccagctca tggccacggc cacactcaag aggaactctg gcacctgcct 4320
gtttgtggcc gtgaaacggc tgatcctttg ctatgagatc cagagaacga agccattcca 4380
cagaaagttc aatgagattg tggctcccgg cagcgtgcag tgcctggcgg tgctcaggga 4440
caggctctgt gtgggctacc cttctgggtt ctgcctgctg agcatccagg gggacgggca 4500
gcctctaaac ctggtaaatc ccaatgaccc ctcgcttgcg ttcctctcac aacagtcttt 4560
tgatgccctt tgtgctgtgg agctcgaaag cgaggagtac ctgctttgct tcagccacat 4620
gggactgtac gtggacccgc aaggccggag ggcacgcgcg caggagctca tgtggcctgc 4680
ggctcctgtc gcctgtagtt gcagccccac ccacgtcacg gtgtacagcg agtatggcgt 4740
ggacgtcttt gatgtgcgca ccatggagtg ggtgcagacc atcggcctgc ggaggataag 4800
gcccctgaac tctgaaggca ccctcaacct cctcaactgc gagcctccac gcttgatcta 4860
cttcaagagc aagttctcgg gagcggttct caacgtgccg gacacctccg acaacagcaa 4920
gaagcagatg ctgcgcacca ggagcaaaag gcggttcgtc ttcaaggtcc cagaggaaga 4980
gagactgcag cagaggcgag agatgcttag agacccagaa ttgagatcca aaatgatatc 5040
caacccaacc aacttcaacc acgtggccca catgggccca ggcgacggca tgcaggtgct 5100
catggacctg cctctgagtg ctgtgccccc ctcccaggag gaaaggccgg gccccgctcc 5160
caccaacctg gctcgccagc ctccatccag gaacaagccc tacatctcgt ggccctcatc 5220
aggtggatcg gagcctagcg tgactgtgcc tctgagaagt atgtctgatc cagaccagga 5280
ctttgacaaa gagcctgatt cggactccac caaacactca actccatcga atagctccaa 5340
ccccagcggc ccaccgagcc ccaactcccc ccacaggagc cagctccccc tcgaaggcct 5400
ggagcagccg gcctgtgaca cctgaagccg ccagctcgcc acaggggcca gggagctgga 5460
gatggcctcc agcgtcagtg ccaagactga gcgggccctc cagtgttgtc caaggaaatg 5520
tagaatcact ttgtagatat ggagatgaag aagacaaatc tttattataa tattgatcag 5580
ttttatgccg cattgttcgt ggcagtagac cacatctgtt cgtctgcaca gctgtgaggc 5640
gatgctgttc catctgcaca tgaaggaccc ccatacagcc tgtctcccac ccctgacaac 5700
ccgagagggc atatggggcc ctgccaacac cacttcctca gcagaaaccc gtcatgacgc 5760
ggctgcttcg gaagcagaca tctggggaca cagcctcagt acccagtctt ttccctagtt 5820
cctgaaactt tcctaggacc ttaagagaat agtaggaggt cctatagcat tcccagtgtc 5880
actagaattt tgaagacagg aaagtggagg ttagtctgtg gccttttttt catttagcca 5940
ttgcacagtc agctgcagaa gtcctgctga ccacctagtc atggacaaag gcccaggacc 6000
agtgacaccc tgcgtccctg tgtgcgttaa gttcattctg ggtcgcagcc atgaagtgtc 6060
accagtatct actactgtga agtcagctgt gctgttttcc attcgcttcc acggcttctg 6120
cctcctgcca taaaaccagc gagtgtcgtg gtgcaggcag gccctgtggc ctgctgggct 6180
gagggaagtc agagccccag ggcgccacga agcagccact gggatacccc accccgcccc 6240
gccnncccce ccccccccnc cagtcnagnc ccgaaatgga gcccccgtga ttagtagccc 6300
gtatgatcac gtagacccac ccaacacact cctgcacact ggccccggcc cacggcacag 6360
caatcccctg cgcgtggatt tcacctcacc ctttgtacca gatgttgagt gaccagctct 6420
gtggccctgt gtcgtcagag gcttgtgatt aactgtggcg gcagacacag cttgtccaca 6480
gcttgggcca ggcttccgct gtcctcccac cggtcggctg cttggcaagg ctgttcagga 6540
cgtgcacttc cccaagtcgg cactgagtgg cccagcacca cctagccctg ccaccccact 6600
gccctcctgg gccttctgct ggatgggcac ctgggggttt ctggtttttt actttttaa 6660
tgtaagtctc agtctttgta attaattatt gaattgtgag aacatttttg aacaatttac 6720
ctgtcaataa agcagaagac ggcagtttta aagttaaaaa aaaaaaaaaa aaaaaaaaaa 6780

<210> 6

<211> 2380

<212> DNA

<213> Homo sapiens

<400> 6
gagcggccgc ccgggcaggt ctagaattca gcggccgctg aattctaggt gctgccggag 60
acctcagggc cccttaaaga ggacccattc ccctgtagac cagtctctgt cccctgcaag 120
ctttacctgc attcttgccc atggcgcttc ccattttctg tgtgatctta tctccctcac 180
ttaactctct gttcctgtgt cttcattcta tgagctggac tgaggccttg ggtgggaaag 240
tgggctctgt attctattcc gtgcctaacc cagcgcctcc ttcttgtgtc tttctccctc 300

```
tctagcctat ctggtcagtc aggcaaccga tcttcctcag gatcattgat ctctgtacct    360
ccaggggcag tgaaccttcc tttccctggg ataatcctca aggctcactg atcaaacctt    420
tgggcttggt tcacaggtta ggtctatgtc agtacgcgac atcagatatt tgtgttcgtc    480
agggttttct aggggaaaag agctggtaga atggaaaagt ggagatttat taggctgcag    540
tctgctagtc caccaatggc tggtagttct tttggaaatg atttatttcc atcccttatg    600
tgtatgagta cctttggcct tcctgtgtct gtgccccatg tgccgtggag cgtggtcgcc    660
acctcctcat cctgatctct ttagggagac acgactctgc caagcccttc ctgccttcaa    720
tgtcagtacc cgcttgactt tccccagtgt tccttcggcg ttggcggaga gatgttgaag    780
ttcaagtatg gtgtgcggaa cccgccggag gccagtgcct ccgagcccat tgccagtcgg    840
gcctccaggc taaatctctt cttccagggg aaaccgcccc tcatgactca acagcagatg    900
tctgctcttt cccgggaagg gatgctagac gccctcttcg ctctctttga agagtgcagc    960
caacccgccc tgatgaagat gaagcacgtg agcagctttg tccagaagta ttccgacacc   1020
atagccgagt tgcgggagct gcagccgtcg gcgagagact tcgaagttcg aagccttgtg   1080
ggctgtggtc acttcgctga agtgcaggtg gttagagaga aggcgaccgg ggacgtctat   1140
gccatgaaaa tcatgaagaa gaaggctttg ctggcccagg aacaggtttc atttttcgag   1200
gaggagagga acatattatc tcggagcacg agtccttgga tcccccagtt acagtacgcc   1260
tttcaggaca aaaataacct ttacctggtc atggaatatc agcctggagg ggatttcctg   1320
tcgcttctga acagatacga ggaccaatta gatgagagca tgatccagtt ttaccttgct   1380
gagctgattt tggctgtcca cagcgtgcac cagatgggat atgtgcatcg agacatcaag   1440
cccgagaaca tcctcatcga ccggacggga gagatcaagc tggtggattt tggatcagcc   1500
gctaagatga attcaaataa ggtggatgcc aaactcccca ttgggacccc ggattacatg   1560
gctccggaag tgttgaccgt gatgaacgag gaccgaaggg gcacatacgg cttggactgt   1620
gactggtggt ctgtcggagt tgttgcttat gagatggttt atgggaagac cccattcaca   1680
gagggaacct ccgcccggac cttcaacaac atcatgaact tccagcggtt tttgaagttc   1740
ccagatgacc ccaaagttag cagtgagctc cttgatctgc ttcagagtct gctgtgtgtc   1800
cagaaagaga gactgaagtt cgagggtctc tgctgccacc ctttctttgc cagaacggac   1860
tggaacaaca tccgtaactc tcctcccccc ttcgtcccca ccctcaagtc tgacgatgac   1920
acctccaatt ttgatgaacc agagaagaat tcgtgggctt tcatcctctg tgtgccagct   1980
gagcccctcg cgttctcagg cgaagagctg ccgtttgtgg gattttcgta cagcaaggca   2040
ctggggtatc ttggtagatc tgagtctgtc gtgtcgagtc tggactcccc tgccaaggtt   2100
agctccatgg aaaagaaact tctcatcaaa agcaaagagc tccaagactc ccaggacaag   2160
tgtcacaagg tatctatctc cacagccggc ctccgtcctt gctccaggat cctccagtca   2220
atatatgccg agggatctgc cgggggccac tgctgagcgg tggtgccgcc tccctcgctg   2280
aagtcgtgcc tccagcagct cagagggaga ggactccagg cccagacatt gccataaatc   2340
ctttaaatct taaccagagg aggccctgga tttaaaaaaa                         2380
```

<210> 7

<211> 631

<212> PRT

<213> Homo sapiens

<400> 7

```
Met Ser Ala Glu Val Arg Leu Arg Gln Leu Gln Gln Leu Val Leu Asp
1               5                  10                  15
Pro Gly Phe Leu Gly Leu Glu Pro Leu Leu Asp Leu Leu Leu Gly Val
            20                  25                  30
His Gln Glu Leu Gly Ala Ser His Leu Ala Gln Asp Lys Tyr Val Ala
        35                  40                  45
Asp Phe Leu Gln Trp Val Glu Pro Ile Ala Ala Arg Leu Lys Glu Val
    50                  55                  60
Arg Leu Gln Arg Asp Asp Phe Glu Ile Leu Lys Val Ile Gly Arg Gly
65                  70                  75                  80
Ala Phe Ser Glu Val Ala Val Val Lys Met Lys Gln Thr Gly Gln Val
                85                  90                  95
Tyr Ala Met Lys Ile Met Asn Lys Trp Asp Met Leu Lys Arg Gly Glu
            100                 105                 110
Val Ser Cys Phe Arg Glu Glu Arg Asp Val Leu Val Lys Gly Asp Arg
        115                 120                 125
```

```
Arg Trp Ile Thr Gln Leu His Phe Ala Phe Gln Asp Glu Asn Tyr Leu
    130                 135                 140
Tyr Leu Val Met Glu Tyr Tyr Val Gly Gly Asp Leu Leu Thr Leu Leu
145                 150                 155                 160
Ser Lys Phe Gly Glu Arg Ile Pro Ala Glu Met Ala Arg Phe Tyr Leu
                165                 170                 175
Ala Glu Ile Val Met Ala Ile Asp Ser Val His Arg Leu Gly Tyr Val
            180                 185                 190
His Arg Asp Ile Lys Pro Asp Asn Ile Leu Leu Asp Arg Cys Gly His
        195                 200                 205
Ile Arg Leu Ala Asp Phe Gly Ser Cys Leu Lys Leu Gln Pro Asp Gly
    210                 215                 220
Met Val Arg Ser Leu Val Ala Val Gly Thr Pro Asp Tyr Leu Ser Pro
225                 230                 235                 240
Glu Ile Leu Gln Ala Val Gly Gly Gly Pro Gly Ala Gly Ser Tyr Gly
                245                 250                 255
Pro Glu Cys Asp Trp Trp Ala Leu Gly Val Phe Ala Tyr Glu Met Phe
            260                 265                 270
Tyr Gly Gln Thr Pro Phe Tyr Ala Asp Ser Thr Ala Glu Thr Tyr Ala
        275                 280                 285
Lys Ile Val His Tyr Arg Glu His Leu Ser Leu Pro Leu Ala Asp Thr
    290                 295                 300
Val Val Pro Glu Glu Ala Gln Asp Leu Ile Arg Gly Leu Leu Cys Pro
305                 310                 315                 320
Ala Glu Ile Arg Leu Gly Arg Gly Gly Ala Gly Asp Phe Gln Lys His
                325                 330                 335
Pro Phe Phe Phe Gly Leu Asp Trp Glu Gly Leu Arg Asp Ser Val Pro
            340                 345                 350
Pro Phe Thr Pro Asp Phe Glu Gly Ala Thr Asp Thr Cys Asn Phe Asp
        355                 360                 365
Val Val Glu Asp Arg Leu Thr Ala Met Val Ser Gly Gly Gly Glu Thr
    370                 375                 380
Leu Ser Asp Met Gln Glu Asp Met Pro Leu Gly Val Arg Leu Pro Phe
385                 390                 395                 400
Val Gly Tyr Ser Tyr Cys Cys Met Ala Phe Arg Asp Asn Gln Val Pro
                405                 410                 415
Asp Pro Thr Pro Met Glu Leu Glu Ala Leu Gln Leu Pro Val Ser Asp
            420                 425                 430
Leu Gln Gly Leu Asp Leu Gln Pro Pro Val Ser Pro Pro Asp Gln Val
        435                 440                 445
Ala Glu Glu Ala Asp Leu Val Ala Val Pro Ala Pro Val Ala Glu Ala
    450                 455                 460
Glu Thr Thr Val Thr Leu Gln Gln Leu Gln Glu Ala Leu Glu Glu Glu
465                 470                 475                 480
Val Leu Thr Arg Gln Ser Leu Ser Arg Glu Leu Glu Ala Ile Arg Thr
                485                 490                 495
Ala Asn Gln Asn Phe Ser Ser Gln Leu Gln Glu Ala Glu Val Arg Asn
            500                 505                 510
Arg Asp Leu Glu Ala His Val Arg Gln Leu Gln Glu Arg Met Glu Met
        515                 520                 525
Leu Gln Ala Pro Gly Ala Ala Ala Ile Thr Gly Val Pro Ser Pro Arg
    530                 535                 540
Ala Thr Asp Pro Pro Ser His Leu Asp Gly Pro Pro Ala Val Ala Val
545                 550                 555                 560
Gly Gln Cys Pro Leu Val Gly Pro Gly Pro Met His Arg Arg His Leu
                565                 570                 575
Leu Leu Pro Ala Arg Ile Pro Arg Pro Gly Leu Ser Glu Ala Arg Cys
            580                 585                 590
Leu Leu Leu Phe Ala Ala Ala Leu Ala Ala Ala Ala Thr Leu Gly Cys
        595                 600                 605
Thr Gly Leu Val Ala Tyr Thr Gly Gly Leu Thr Pro Val Trp Cys Phe
    610                 615                 620
```

Pro Gly Ala Thr Phe Ala Pro
625 630

<210> 8

<211> 1765

<212> DNA

<213> Homo sapiens

<400> 8
atgttgaagt tcaaatatgg agcgcggaat cctttggatg ctggtgctgc tgaacccatt 60
gccagccggg cctccaggct gaatctgttc ttccagggga aaccaccctt tatgactcaa 120
cagcagatgt ctcctctttc ccgagaaggg atattagatg ccctctttgt tctctttgaa 180
gaatgcagtc agcctgctct gatgaagatt aagcacgtga gcaactttgt ccggaagtat 240
tccgacacca tagctgagtt acaggagctc cagccttcgg caaaggactt cgaagtcaga 300
agtcttgtag gttgtggtca ctttgctgaa gtgcaggtgg taagagagaa agcaaccggg 360
gacatctatg ctatgaaagt gatgaagaag aaggctttat tggcccagga gcaggtttca 420
ttttttgagg aagagcggaa catattatct cgaagcacaa gcccgtggat cccccaatta 480
cagtatgcct ttcaggacaa aaatcacctt tatctggtca tggaatatca gcctggaggg 540
gacttgctgt cacttttgaa tagatatgag gaccagttag atgaaaacct gatacagttt 600
tacctagctg agctgatttt ggctgttcac agcgttcatc tgatgggata cgtgcatcga 660
gacatcaagc ctgagaacat tctcgttgac cgcacaggac acatcaagct ggtggatttt 720
ggatctgccg cgaaaatgaa ttcaaacaag atggtgaatg ccaaactccc gattgggacc 780
ccagattaca tggctcctga agtgctgact gtgatgaacg gggatggaaa aggcacctac 840
ggcctggact gtgactggtg gtcagtgggc gtgattgcct atgagatgat ttatgggaga 900
tccccttcg cagagggaac ctctgccaga accttcaata acattatgaa tttccagcgg 960
tttttgaaat ttccagatga ccccaaagtg agcagtgact ttcttgatct gattcaaagc 1020
ttgttgtgcg gccagaaaga gagactgaag tttgaaggtc tttgctgcca tcctttcttc 1080
tctaaaattg actggaacaa cattcgtaac tctcctcccc ccttcgttcc caccctcaag 1140
tctgacgatg acacctccaa ttttgatgaa ccagagaaga attcgtgggt ttcatcctct 1200
ccgtgccagc tgagcccctc aggcttctcg ggtgaagaac tgccgtttgt ggggttttcg 1260
tacagcaagg cactggggat tcttggtaga tctgagtctg ttgtgtcggg tctggactcc 1320
cctgccaaga ctagctccat ggaaaagaaa cttctcatca aaagcaaaga gctacaagac 1380
tctcaggaca agtgtcacaa ggtatttatt tccgcagccg gcctccttcc ttgctccagg 1440
atcctcccgt ccgtatatgc caagggatcc gcccggggcc gctgctggct ctgagccgcc 1500
tgatccgtag agagtgaggc gctcctgcct tcgctgaagt cgcgcctcca gcagctcaga 1560
gggagatgaa ttcgggcctt gctgttgctg taaatccttt aaatctaaac cagaggaggc 1620
cctggattta aacagtccgt ttctcagcat gacccagcca gatgtctgct tcttccggca 1680
ggtggcctgg gtcctcacct gtggctgaga tacatcccat ctgctttgag tgatgcgaag 1740
tctctcttcc tagtctttta aaact 1765

<210> 9

<211> 497

<212> PRT

<213> Homo sapiens

<400> 9
Met Leu Lys Phe Lys Tyr Gly Ala Arg Asn Pro Leu Asp Ala Gly Ala
1 5 10 15
Ala Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
20 25 30
Gly Lys Pro Pro Phe Met Thr Gln Gln Gln Met Ser Pro Leu Ser Arg
35 40 45
Glu Gly Ile Leu Asp Ala Leu Phe Val Leu Phe Glu Glu Cys Ser Gln
50 55 60

```
Pro Ala Leu Met Lys Ile Lys His Val Ser Asn Phe Val Arg Lys Tyr
65                  70                  75                  80
Ser Asp Thr Ile Ala Glu Leu Gln Glu Leu Gln Pro Ser Ala Lys Asp
                85                  90                  95
Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
            100                 105                 110
Val Val Arg Glu Lys Ala Thr Gly Asp Ile Tyr Ala Met Lys Val Met
        115                 120                 125
Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
    130                 135                 140
Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
145                 150                 155                 160
Gln Tyr Ala Phe Gln Asp Lys Asn His Leu Tyr Leu Val Met Glu Tyr
                165                 170                 175
Gln Pro Gly Gly Asp Leu Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
            180                 185                 190
Leu Asp Glu Asn Leu Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
        195                 200                 205
Val His Ser Val His Leu Met Gly Tyr Val His Arg Asp Ile Lys Pro
    210                 215                 220
Glu Asn Ile Leu Val Asp Arg Thr Gly His Ile Lys Leu Val Asp Phe
225                 230                 235                 240
Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Met Val Asn Ala Lys Leu
                245                 250                 255
Pro Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met
            260                 265                 270
Asn Gly Asp Gly Lys Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser
        275                 280                 285
Val Gly Val Ile Ala Tyr Glu Met Ile Tyr Gly Arg Ser Pro Phe Ala
    290                 295                 300
Glu Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg
305                 310                 315                 320
Phe Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Asp Phe Leu Asp
                325                 330                 335
Leu Ile Gln Ser Leu Leu Cys Gly Gln Lys Glu Arg Leu Lys Phe Glu
            340                 345                 350
Gly Leu Cys Cys His Pro Phe Phe Ser Lys Ile Asp Trp Asn Asn Ile
        355                 360                 365
Arg Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp
    370                 375                 380
Thr Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Val Ser Ser Ser
385                 390                 395                 400
Pro Cys Gln Leu Ser Pro Ser Gly Phe Ser Gly Glu Glu Leu Pro Phe
                405                 410                 415
Val Gly Phe Ser Tyr Ser Lys Ala Leu Gly Ile Leu Gly Arg Ser Glu
            420                 425                 430
Ser Val Val Ser Gly Leu Asp Ser Pro Ala Lys Thr Ser Ser Met Glu
        435                 440                 445
Lys Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys
    450                 455                 460
Cys His Lys Val Phe Ile Ser Ala Ala Gly Leu Leu Pro Cys Ser Arg
465                 470                 475                 480
Ile Leu Pro Ser Val Tyr Ala Lys Gly Ser Ala Arg Gly Arg Cys Trp
                485                 490                 495
Leu
```